California Energy Commission Workshop

"Exploring Alternative Wholesale Electricity Market Structures for California"

November 7, 2001 (9:00 am – 4:15 pm)
California Energy Commission Hearing Room A
1516 Ninth Street
Sacramento, California 95814

Facilitated by EPRI
3412 Hillview Avenue
Palo Alto, CA 94304

ANNOUNCEMENT

California Energy Commission Workshop on Exploring Alternative Wholesale Electricity Market Structures for California

November 7, 2001 (9:00 am – 4:15 pm)
California Energy Commission Hearing Room A
1516 Ninth Street
Sacramento, California 95814

The intent of this workshop is to focus on the relationship between market structure and reliability. Given California's recent experience with competitive electricity markets it appears that the design and implementation of the market was inadequate from the perspective of ensuring reliability. In California, the principal market mechanism that was supposed to stimulate construction of new generation was the Power Exchanges' hourly market clearing price for energy. For generators who depended solely on the PX's energy market for revenue, profitability was largely a function of high prices paid during the summer. As a result of this market structure, generators had a strong incentive to allow operating reserves to drop below the level needed to ensure reliable service. Generators can accomplish this by either not building new generation or withholding existing generation from the market. This kind of behavior on the part of generators contributed to the volatile electricity prices in California during the summer of 2000 and winter of 2000-2001. Looking at this issue from a demand-side perspective, the insulation of customers from high electricity prices through a legislated rate freeze, muted any incentive to reduce loads. The failure to incorporate into the design of the market demand responsive load programs also contributed to price volatility and poor reliability.

This workshop will evaluate alternative market structures from two perspectives; first, the ability of alternative market structures to incent the timely addition of new generation so as to reduce price volatility and contribute to reliable service, and second, the ability of various market structures to facilitate robust demand-side participation in the market through programs or mechanisms that allow consumers to respond to real time prices.

To facilitate this discussion, Andy Ford of Washington State University and Stephen Lee of EPRI will present the results of their recent analyses using models they've developed that look at alternative market structures and generation investment patterns. The workshop will also include discussion on the role of the newly created California Consumer Power and Conservation Financing Authority (California Power Authority) and the impact of its investment decisions on other market participants and the future course of electricity markets in the state.

This Workshop is open to the public. For directions to the CEC building, go to the following Internet link: http://www.energy.ca.gov/commission/directions.html

For more information contact Steve Lee, 650/855-2486, slee@epri.com.

California Energy Commission Workshop on
Exploring Alternative Wholesale Electricity Market Structures for California

November 7, 2001 (9:00 am – 4:15 pm)
California Energy Commission Hearing Room A
Sacramento, California

Preliminary Agenda

Time	Item	Presenters or Panelists
9:00	Calling Meeting to Order Introductions and Workshop Objectives	Karen Griffin (CEC)
9:15	Propensity of a Competitive Power Market towards Boom/Bust Cycles – Theory and Insights	Andy Ford (WSU)
10:00	Questions and Answers	
10:30	Comparison of a Competitive Wholesale Power Market with Alternative Structures through a Long Term Power Market Simulation Model	Stephen Lee (EPRI)
11:15	Questions and Answers	
11:30	Panel Discussion on the Long Term Objectives of the California Wholesale Power Market	Moderator: Karen Griffin, (CEC) Panelists: Kellan Fluckiger (CPCFA), Eric Woychik (Strategy Integration), CPUC / Legislative Staff*
12:30 pm	Lunch Break	
1:30 pm	Panel Discussion on the Means of Reducing Price Volatility in a Competitive Power Market from the Supply and the Demand Sides	Moderators: Stephen Lee, Michael Jaske (CEC) Panelists: John Chandley (LECG), Carl Blumstein (UCEI), Steven Stoft
2:45 pm	Panel Discussion on the Alternative Wholesale Market Structures for California	Moderator: Andy Ford Panelists: Philippe Auclair (Mirant), Charles Kolstad (UCSB), Lorenzo Kristov (ISO)
4:00 pm	Summary and Concluding Remarks	Karen Griffin
4:15	Workshop Adjourns	

* to be announced.

List of Issues and Questions

Panel Discussion on the Long Term Objectives of the California Wholesale Power Market

What are the long term objectives of the California wholesale power market?

What lessons have we learned from the California power crisis of 2000-2001 about the importance of electricity market structures?

Are there other competitive wholesale electricity markets outside of California that are functionally competitive, or have these markets simply benefited from there being a surplus of capacity present when they were initiated? Lessons from the UK, PJM, New England, New York, others.

What are the structural shortcomings of a purely competitive wholesale power market?

Panel Discussion on the Means of Reducing Price Volatility in a Competitive Power Market from the Supply and the Demand Sides

Is price volatility an inevitable component of a competitive wholesale power market? Is the volatility of a daily, seasonal or cyclical nature?

What is the level of price volatility that retail end-users can tolerate?

How does the wholesale market structure affect investment patterns (i.e. boom/bust cycles) and impact prices and price volatility?

What other factors, unrelated to market structure, contribute to price volatility and boom-bust cycles? How can these be mitigated?

How does the wholesale market structure affect the ability of market participants to recover their investments in new generation and manage risk? What are the sources of risk?

Can price caps be used as an effective means of controlling price volatility without having a detrimental impact on investment in new generation?

Would providing capacity payments to generators be a more effective means than price caps of limiting price volatility and generator and investor risk?

What end-use loads should be targeted for participation in demand responsiveness programs and mechanisms? How can technology, which provides market price signals to end-users and automatically controls some loads in response to those signals, facilitate demand-side participation in the market?

What actions by the California Public Utilities Commission and public utility governing boards are needed to facilitate the demand-side market?

- Are real-time pricing tariffs enough, or is there still a role for load curtailment programs activated by system operating reserve conditions?

Is it possible to assess how much demand responsive load, actively engaged in market price setting through real-time pricing or other demand responsive programs, is required to limit price volatility to an acceptable level?

How does market design influence the success or failure of greater demand-side participation in the market?

Panel Discussion on the Alternative Wholesale Market Structures for California

What are some of the possible wholesale market structure options for California and what are the pros and cons of each in terms providing transparent pricing and ensuring system reliability?

- Limited bilateral market with robust spot market
- Robust bilateral market with limited spot market
- Hybrid market with both nonregulated and governmental participants such as the Power Authority
- Others

How could capacity payments be structured so as to provide appropriate pricing signals that new generation is needed?

What roles are most publicly beneficial for *state investment* to achieve the objectives of a robust and reliable power market?

What are reasonable limitations on the roles of *state investment* such that the private investors continue to have sufficient financial incentive to build future generating capacity to meet the State's need for power?

Short Bios for the 11/7/2001 CEC Workshop

Philippe Auclair

Philippe Auclair is Manager of Regulatory Affairs for Mirant, Americas Inc. He is responsible for market design issues in the Pacific Northwest for both power and natural gas, and in Alberta and British Columbia for power issues only. Key areas of emphasis include the RTO West formation, the Alberta power market design study, and Northwest pipeline issues. Before joining Mirant, Mr. Auclair held positions at the California Electricity Oversight Board (EOB), California Public Utilities Commission (CPUC) and the California Energy Commission (CEC). While at the EOB, Mr. Auclair was appointed to the CAISO Market Advisory Group. While at the CPUC, he was an advisory member of the Board of Governors of the California Power Exchange. His interests include electricity spot and financial contract market designs with a focus on congestion management pricing issues. Mr. Auclair holds an M.A. in economics from U.C. Davis and a B.S. from the State University of New York, Albany.

Carl Blumstein

Carl Blumstein is an energy policy analyst at the University of California Energy Institute. His research interests center on energy efficiency and energy policy. Recent work has included publications concerning restructuring in the electric power supply industry, the Strategic Petroleum Reserve, and the role of market transformation in energy efficiency programs. Dr. Blumstein is Chairman of the Board of the American Council for an Energy-Efficient Economy (ACEEE), a member of the Board of Governors of the California Power Exchange, and a member of the Board of Trustees of the Consortium for Energy Efficiency. He holds a Ph.D. from the University of California at San Diego.

John D. Chandley

John Chandley is a Principal at LECG and is a member of the electricity market design team led by Harvard's William Hogan and LECG's Scott Harvey. This team has been responsible for designing and writing the market rules for PJM and New York, the market reform proposals for New England and the market rules in Ontario (with Larry Ruff, then at Putnam, Hayes and Bartlett). While advising the Midwest ISO and assisting clients in the Alliance and Southeast RTO regions, the team is also significantly involved in the FERC's efforts to create Regional Transmission Organizations. Toward this end, Mr. Chandley recently prepared a paper on standard market design principles, and this paper will soon be published in the *Electricity Journal*. Prior to joining the Hogan/Harvey market design team, Mr. Chandley was an Assistant Chief Counsel at the California Energy Commission, where he worked for 20 years, and was involved in the early restructuring efforts in California.

Short Bios for the 11/7/2001 CEC Workshop

Kellan Fluckiger

Kellan Fluckiger brings a substantial and diverse career history to his current position as senior advisor to the Chair and CEO of the newly created California Power Authority. At the Power Authority he functions as a policy advisor in all areas of the Power Authority's responsibilities, including interface with other state and federal agencies, investment planning, market design, and other business activities. He served as an energy advisor to the Davis administration for six months prior to assuming his current position in September of 2001. Before that, he was with the California ISO from its inception in 1997 to March of 2001 where he served in various leadership positions, including Chief Operations Officer. A native Californian, Mr. Fluckiger's energy career began at Pacific Gas & Electric in 1977 before he moved to Phoenix to join Arizona Public Service Company where he spent 13 years in system operations. He went on to become Senior Manager of Network Operations for the Idaho Power Company prior to becoming a member of the Cal-ISO Start-Up Team. Mr. Fluckiger is a graduate of Ottawa University.

Andrew Ford

Andrew Ford has worked on electric industry problems for 25 years. His dissertation used computer simulation to understand the financial problems of the nation's investor-owned utilities. His doctoral degree is from the Public Policy and Technology Program at Dartmouth College in 1975. Dr. Ford worked in the Energy Policy Group at the Los Alamos National Laboratory in the 1970s and 1980s. His research focused on power plant development in the Rocky Mountain West and the improvement of the policy models used by the U.S. Department of Energy. While teaching systems analysis and simulation at USC, his research focused on incentives for efficiency investments in the Northwest and electric vehicle purchases in California. Dr. Ford is now Professor of Environmental Science and Regional Planning at Washington State University where he teaches modeling with an emphasis on environmental problems. Dr. Ford's recent research concentrates on the transition from regulation to competition in the electric industry. It focuses on the use of interactive simulation to understand the implications of boom and bust in the western electric system.

Short Bios for the 11/7/2001 CEC Workshop

Karen L. Griffin

Karen Griffin manages the Electricity Analysis Office of the California Energy Commission. She is responsible for the Commission's policy analysis of the restructured electricity industry and the supporting technical studies of supply adequacy in the western US and Canada. Ms. Griffin participated in the cross-industry teams that developed the new industry structure, served on the Independent System Operator's steering committee for its reliability study, and participates in long-term grid planning. She has managed electricity functions at the Commission since 1991, and held other management and technical lead positions at the Commission for the previous decade. Before joining the Energy Commission, Ms. Griffin served as an economist and program analyst at the US Department of Energy, the US Department of Commerce, the US Department of the Treasury, and the Norfolk, VA Department of Redevelopment and Public Housing. Ms. Griffin received her undergraduate degree from Duke University and a M.P.A. from Cornell University.

Michael R. Jaske

Dr. Michael Jaske currently acts as a senior policy analyst in the Executive Office of the California Energy Commission (CEC) as a member of the Strategic Issues Integration Group. For 20 years, he was the Chief Demand Forecaster giving technical direction for the Commission Staff's independent demand forecast. Dr. Jaske plays an active role in the development and advocacy of the CEC's positions on retail market structure. Dr. Jaske has been involved in California's electricity restructuring process from the outset, focusing on retail market structure issues, direct access, unbundling utility services, competitive provision of revenue cycle services, collection and communication of load data to support scheduling and settlement, and improving demand responsiveness in electricity markets. Dr. Jaske is a member of the IEEE Power Engineering Society. Dr. Jaske serves on the Energy Policy Committee of IEEE-USA to educate national policymakers on electricity issues. Dr. Jaske's educational background includes a BS in Chemical Engineering from Oregon State University, and a MS and Ph.D. in Systems Science, both from Michigan State University.

Short Bios for the 11/7/2001 CEC Workshop

Charles Kolstad

Charles Kolstad is the Donald Bren Professor of Environmental Economics and Policy at UC Santa Barbara. He received his PhD from Stanford (1982), his MA from Rochester and his BS from Bates College. Most of Prof. Kolstad's research has been in the area of regulation, particularly environmental regulation. Recently, he has also done work in environmental valuation theory. He is particularly interested in the role of information in environmental decision-making and regulation. Currently he has a major research project on the role of uncertainty and learning in controlling the precursors of climate change. His past work in energy markets has focused on coal and electricity markets, including the effect of air pollution regulation on these markets. Prof. Kolstad is the editor of Resource and Energy Economics and the president of the Association of Environmental and Resource Economists (AERE). He is a member of the USEPA's Clean Air Act Compliance Analysis Committee and has served on numerous other advisory boards, including the Environmental Economics Advisory Committee of the USEPA Science Advisory Board.

Lorenzo Kristov

Lorenzo Kristov is the Manager of Market Design in the Policy Office of the California ISO. The Market Design group analyzes and assesses how the design features of the ISO's markets (for ancillary services, congestion management and real-time imbalance energy) affect their performance and efficiency. In this capacity, Dr. Kristov examines the market rules, procedures, and system designs to identify inefficiencies and opportunities for gaming and exercise of market power. He and his staff develop improvements to market design features and market rules to enhance market efficiency and promote workable competition. Prior to joining the ISO, Dr. Kristov was an Energy Economist at the California Energy Commission, and represented the Commission in retail electric restructuring proceedings and working groups at the California Public Utilities Commission. In this position he worked extensively on retail market design issues, including settlement data quality and the unbundling of revenue cycle services, and helped develop rules and protocols for processing and exchanging end-use meter data. Prior to joining the California restructuring process, Dr. Kristov lived in Southeast Asia and worked in the area of private power development. He received his Ph.D. in Economics from the University of California at Davis, specializing in international macroeconomics and monetary theory.

Short Bios for the 11/7/2001 CEC Workshop

Stephen T. Lee

Dr. Stephen T. Lee is Area Manager, Grid Operations and Planning, in EPRI. Dr. Lee has over 30 years of power industry experience. He received his S.B., S.M. and Ph.D. degrees from M.I.T. in Electrical Engineering, majoring in Power System Engineering. He worked for Stone & Webster Engineering in Boston, Systems Control, Inc. (now ABB) in Palo Alto, and he was Vice President of Consulting, Western Region for Energy Management Associates, Inc. Before joining EPRI in 1998, Dr. Lee was an independent Consultant in utility planning and operation. At EPRI, Stephen Lee is responsible for managing research programs for grid operations and planning. He is also active in cooperative projects with the North American Electric Reliability Council (NERC). He is closely involved in the latest developments of transactions management, wide-area security applications and congestion management in North America. He is also the project manager of the EPRI Transmission Reliability Initiative. In recent months, he has been conducting research in long-term power market simulation and transmission planning for a competitive market.

Steven Stoft

Steve Stoft received a B.S. in engineering math and a Ph.D. in economics from U.C. Berkeley. He has worked for LBNL, the U.C. Energy Institute and FERC. He presently consults for PJM, the California EOB and the Alberta Trailer Company. IEEE/Wiley will be publishing his book, "Power System Economics: Designing Markets for Electricity" early next year. Much of it is currently available at www.stoft.com.

Eric Charles Woychik

Eric Charles Woychik, the Principal and founder of Strategy Integration Inc., is a consultant on electric and gas market structure and policy. He has developed a wide range of strategies for consumers, government, regulatory bodies, and other stakeholders located in California, the U.S., and other countries. Mr. Woychik received a B.S. in Environmental Planning & Policy Analysis from University of California Davis and a M.S. in Economics from New Mexico State University. Prior to 1992, Mr. Woychik was a Principal at Synergic Resources Corporation, Commissioner Advisor and Staff Member of the California Public Utilities Commission, Staff Member of the California Energy Commission, and consultant for Davis Alternative Technology Associates and DOE (Lawrence Berkeley Laboratory). For the first part of two decades, much of this work involved electric and gas industry reform in California and the U.S. Since 1984, Mr. Woychik has worked with 16 countries and advised for the governments of Australia, Canada, Ghana, Indonesia, Kazakhstan, Mexico, Norway, Poland, Russia, Sweden, Ukraine, and a number of states in the U.S., on energy market structure. In 1998 he was seated by the Federal Energy Regulatory Commission to the Board of the California ISO and in 2000 was appointed to the CAISO Market Advisory Group.

Simulation Scenarios for the Western Electricity Market

A Discussion Paper for the California Energy Commission Workshop on
Alternative Market Structures for California

November 2001

Professor Andrew Ford
Program in Environmental Science and Regional Planning
Washington State University

Summary

This paper was written to promote discussion at a workshop on alternative electricity markets in California. Workshop participants will discuss changes on both the supply side and the demand side of the wholesale market. On the supply side, participants will comment on incentives to promote more timely investments by private developers. They will also discuss the appropriate role of the California Power Authority. On the demand side, participants will discuss ways to encourage electricity consumers to participate more actively in wholesale markets, and they will comment on the impact of the legislated freeze in retail rates.

This paper argues that competitive electricity markets are prone to the same cycles of boom and bust that appear in commodity markets and in specialized industries like real estate. The paper presents simulation scenarios of how boom and bust could appear in the western electricity system. The "business as usual" scenario envisions a system at the crest of a building boom and on the verge of a bust in wholesale prices. The system would then experience a lull in construction before the next building boom. Without fundamental changes in the wholesale markets, the next construction boom would come too late to prevent a decline in reserve margins and the reappearance of price spikes. If we continue with the current market structure, we run the risk of exposing the western electricity markets to another round of reliability alerts and price spikes.

This paper argues that western markets could be improved by capacity payments to promote more timely private investment in new power plants. The Power Authority could also make the needed investments, but it must be prepared for a large and permanent commitment.

On the demand side, the paper argues that removing the legislative freeze on retail rates does not lead to long-term improvement in wholesale market performance. The more effective approach on the demand side is to implement programs to allow selected customers to respond to wholesale prices in real-time.

Contents

Introduction

The blueprint for a competitive electric industry in California was issued in 1994 and implemented by the Legislature in 1996. The new markets opened for business in 1998. By the summer of 2000, a full-blown crisis had emerged in the form of unprecedented outages and price spikes. The crisis conditions continued through the fall of 2000, spread throughout the west, and continued into the winter and spring of 2001. Then, to the surprise of many, chronic outages and price spikes did not appear in the summer of 2001. New power plants came on line, and many more entered construction. As the year 2001 draws to a close some are predicting that the current building boom will lead to a glut of electricity supply. It appears that the western electric system is experiencing the boom and bust pattern that has appeared in other industries.

Industries with Boom and Bust

Many industries have experienced persistent cycles of boom and bust. The cyclical tendency is especially strong in the commodities. A commodity is usually defined as an undifferentiated product, often supplied by many small, independent producers. Examples include mineral products (i.e. aluminum and copper), forest products (i.e. lumber and pulp) and agricultural products (i.e. coffee and cattle). The instability in the commodity industries is costly for each industry, for its customers and for the nations that depend on commodity exports for the bulk of their hard currency (Sterman 2000).

The commodity industries suffer from chronic instability despite the fact that their products may be stored in inventory as a buffer between production and consumption. Buffer stocks do not exist in the electricity industry because electricity cannot be stored (except at great cost). Lacking buffer stocks, the electric industry looks to extra generating capacity to absorb the variations in supply and demand. In this sense, the electric industry is similar to the real estate industry with the reserve margin in the electric industry corresponding to the vacancy rate in the real estate industry. The industries are similar in several other respects as well. Developers in both industries confront significant delays for permitting and construction. Also, both industries are capital intensive, so developers face the challenge of recovering high fixed costs. Real estate developers strive to recover their investment by renting space in a competitive market; power plant developers aim to recoup their investment by selling electric energy into a competitive market. But the two industries are quite different in age. We have less than a decade of experience with restructured electricity markets in the USA, but real estate has a long history of competitive markets.

The Real Estate Construction Cycle

The long history of real estate is dominated by a series of exuberant building booms and subsequent busts. To illustrate, Figure 1 shows the pattern of boom and bust documented in Homer Hoyt's detailed account of land values in Chicago. The chart shows land values, new construction and business activity, all scaled in percent variation from a normal value. Land values show the greatest variation. For example, land was valued at 80% below normal in 1830; reached normal values by 1835; then shot off the chart. Within a few years, land values were 80% below normal. This was the first of five booms in land values over the one hundred year period. Hoyt's data on new construction begins in 1854, and Figure 1 shows four construction booms in construction over the 80 years. The final boom in the 1920s was the most sustained with unusually high construction continuing in 1926-1928 even though land values were declining.

Hoyt described surges in population as an important external factor, but the key to the boom-bust pattern was the way investors reacted to the population surges. In a typical example, developers did not react in time to prevent land values from increasing far beyond the increase in population. The high prices then led to an exuberant response, described by Hoyt (1933, p 387) as follows:

Developers scramble to build at many locations around the city, and a great many men work secretly and independently on a great variety of structures in many sections of the city. There is no central clearing *house to correlate the impending supply of buildings with the probable demand, so that when all these plans came to fruition, an astonishing number of new structures had been erected.*

This overreaction sets the stage for the bust: "Gross rents fall, and net rents fall even faster. Land values plummet, and foreclosures are everywhere."

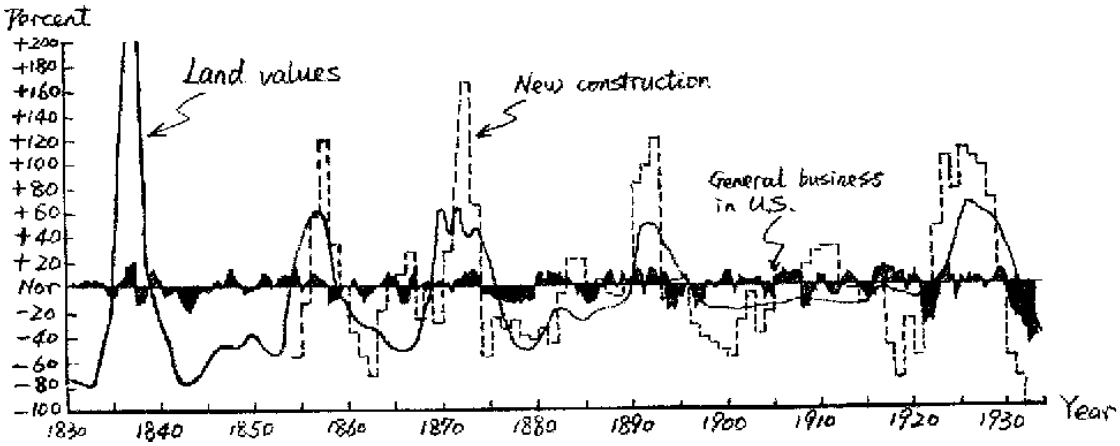


Figure 1. Land values and construction cycles in Chicago.
(traced from Homer Hoyt's *One Hundred Years of Land Values in Chicago.)*

Hoyt concluded his study by speculating that the "real estate cycle itself may be a phenomenon that is confined chiefly to young or rapidly growing cities." But population surges are just one of many external factors that might set the stage for a boom and bust in construction. In more fully developed cities, the external factor may take the form of a surge in income, as happened in cities like Dallas and Boston during the 1970s. The Boston housing market was the focus of a modeling study to understand the developers' response to surges in income and employment. For example, an increase in employment caused housing prices to rise and construction to increase. "Prices peak when so much construction occurs that the stock of housing overshoots its target; prices then decline. This process sets off a repeating cycle in prices, construction and housing stock" (DiPasquale 1996).

It is natural to attribute a particular building boom to an external event like a surge in population in Chicago or a surge in income in Boston. But a full understanding requires us to look at internal as well as external factors. Hoyt looked closely at developers' decision-making in Chicago and concluded that there was no way for developers to keep track of the number of buildings that were under construction. When all the buildings were finally completed, "an astonishing number of new structures had been erected." DiPasquale looked at developer's decision-making in Boston through the lens of structural models. The model with the best statistical explanation of housing construction tells a similar story as Chicago – the developers simply built too much housing during the boom and "the stock of housing overshoots the target."

The overbuilding observed in Chicago and Boston arises, in part, from developers ignoring or discounting the future impact of construction "in the pipeline." The tendency to understate the impact of the "pipeline" appears in industries, ranging from "A to Z, from aircraft to zinc" (Sterman 2000, p. 792). In the case of real estate, developers from the 1980s and 1990s concentrated on picking the best location and bringing the project to market. Student interviews summarized by Sterman (2000, p. 702) revealed that developers focused on selecting an attractive building site, winning financial backing and navigating the permitting process. According to one developer:

Location is a bigger factor than the macro market.
I know it's cliché but really the key to real estate is location, location, location.

Unless prompted by the interviewers, none of the developers mentioned cycles in construction or the impact of the construction pipeline. When prompted to comment on construction cycles, developers admitted that they "never looked at cycles" or "really had no sense for cycles." One developer responded :

Quite frankly, I am lousy when it comes to cycles. I think they exist but don't pay a lot of attention to them. There are too many other factors that affect supply and demand. External factors make it difficult to look at cycles. In fact, I think they probably negate them.

The interviews revealed that developers would sometimes use computer models to keep track of vacancies, rents and cash flow. However, these models were not designed for learning about the dynamics of boom and bust. Their real purpose was to build the case for financing:

Market analysis was not done for decision-making; it was done to obtain financing … Frankly, during that period of time [the boom of the 1980s] you were concerned about getting the deal done and didn't really care about cycles – it was all ego and pressure to do the deal.

The interviewers also revealed a tendency toward herding behavior during the peak of the boom. One developer described the psychological pressures as follows:

There is a tremendous pressure to follow the crowd. If you are standing on the sidelines and not making investments and your competitors are collecting fees and placing funds, what are you going to do?

Learning from The Real Estate Construction Cycle

This brief review has focused on real estate developers' decision-making to gain insight as to how construction cycles might unfold in the electric industry. As we think about power plant construction, it's certainly important to account for the physical factors such as capital intensity and construction lead times and to account for the tendency of developers to not fully account for construction "in the pipeline." We should also expect psychological factors to play a role in shaping investor behavior. The key psychological factors that contribute to instability in markets include a focus on external events, a tendency for "herding" and "groupthink," and simple denial that cycles could exist (ERPI 2000).

Although we have learned to live with construction cycles in the real estate industry, it's not at all clear that we should tolerate construction cycles in the power industry. A fundamental difference in the two industries is the flexibility of the demand side. In real estate, we deal with the periods of low vacancies and high rents by adjusting our demand for floor space. When rents are unusually high, we squeeze into smaller quarters and wait for the boom in construction to bring rents back to normal levels. In electric power, however, customers have little ability to react when reserve margins are low and prices are high. (When wholesale prices are high, we can't "squeeze" fewer kilowatts through our appliances.) With the current market structure, our ultimate response to dangerously low reserve margins is to schedule rolling blackouts to protect the integrity of the system.

The extraordinary reliability requirement of the electric industry sets it apart from industries like real estate. The challenge for the workshop participants is to suggest changes in wholesale markets that will allow the electric system to grow without the periods of tight supply that appear during the course of a construction cycle. The west needs new market structures that will avoid a replay of the price spikes and outages that appeared in 2000-2001.

History of the Electric Industry

It's been over one hundred years since Edison invented the incandescent lamp and almost as long since Thomas Insull invented the investor-owned utility (IOU). The IOUs have managed to expand the nation's generating capacity to keep reserve margins at safe levels during times of extraordinary change. The nation has experienced several wars, a major depression, an extreme drought and the IOUs have experimented with the introduction of complex, long-lead time technologies such as nuclear power.

Throughout this period, the IOUs have expanded the generating capacity to provide the margin of reserves needed for reliable service. Their success is due, in large part, to rate of return regulation and their clear obligation to serve the demand within their service territories.

With restructuring, we look to the markets to build the power plants that will be needed in the future. The IOUs obligation to expand capacity has been replaced by competition among suppliers. Some believe this decision was taken without serious debate. According to the San Jose Mercury News (Nov 30, 2000), for example:

California's deregulation effort was based on an unquestioning faith in the power of the free market. Despite official state forecasts that electricity demands would increase, for example, there was virtually no discussion of whether California's generating capacity would keep pace.

The crisis conditions of 2000-2001 have revealed the serious consequences of insufficient power plant construction in the western system. The lag in construction is documented in a recent report by the Northwest Power Planning Council (NPPC 2000) and in Senate Testimony by Steve Oliver (2000). Oliver suggests that the lack of construction might be attributed to numerous uncertainties that surround the transition to competitive markets. But he warned that the problem could be fundamental and persistent:

The two-to-three year time lag in the market's ability to respond to price signals with new generation supplies may reflect an inherent challenge for competitive electricity markets.

Power plant developers around the US have responded to the market signals with a major increase in proposed projects. An EPRI review of proposed power plants in the US "anticipated that approximately 212 GW of new gas-fired capacity additions could appear over the next five years." This would be approximately "two to three times more than would be needed to keep pace with demand growth. The supply-demand balance would be shifted significantly, and market prices would probably fall substantially below the level needed to support new construction." The review concluded that different regions of the country "could move from boom to bust in just a few years" (EPRI 2000). Over time, the topic of boom and bust is appearing more frequently in the news. For example:

- An article in the Wall Street Journal (July 7, 2000) reports that Duke Energy Corp builds cycles of boom and bust into their own scenario analysis of electricity markets around the world. "There is little doubt there will be over-building. It will happen. The question is when?" says James Donnell, President of Duke's North American Energy Unit.

- The New York Times (August 22, 2001) reports on proposals for 350 GW of new capacity that could be on line as early as 2004, said to be enough to boost the nation's capacity by 50%. The article reported that many energy experts expect power plant construction to run in boom-and-bust cycles.

- According to an editorial in the Sacramento Bee (September 2, 2001), the goal of California's new Power Authority is to build enough power plants for a 15% reserve margin, a policy said to provide "anti-blackout insurance" and to "break the cycle of boom and bust."

- A recent article in Public Utilities Fortnightly (September 15, 2001) looks at the "fears of a coming glut in energy supply" to explain the recent loss in equity values in the power generation industry.

These recent articles have appeared around the same time as a major construction boom is occurring in the west. If we look back to the years prior to the building boom, however, we find only a few studies dealing with the prospects for a boom and bust.

Previous Studies of Boom & Bust

I have reviewed a variety of computer modeling approaches that could help one understand the potential for boom and bust in power plant construction. I found a few models that simulate new construction as an internal variable, but these relied on a combination of optimization and "perfect foresight" to calculate construction over time. This approach may appeal to a theory of "rational expectations," but it precludes a serious consideration of boom and bust (ERPI 2000). A new approach was needed, one which represents "decision-making as it is, not as it should be, nor how it would be if people were perfectly rational" (Sterman 2000, p. 597).

I began the development of more realistic models in the summer of 1998. The first model represented the average annual energy loads and resources in the WSCC, the Western System Coordinating Council (Ford 1999). It assumed that power plant developers watch the general trend in spot market prices and extrapolate those trends into the future. The extrapolated prices were compared to the levelized cost of a new CC for purposes of permit applications and construction starts. These assumptions, when combined with the delays for permitting and construction, caused construction to appear in waves of boom and bust. These results were published near the end of 1999.

A few months later, the CEC released a report on <u>Market Clearing Prices Under Alternative Resource Scenarios 2000 – 2010</u> (CEC 2000). By this time, a large number of power plant proposals had accumulated, and staff could see the early signs for a potential boom in construction. They designed a "rapid development" scenario and a "cautious development" scenario for power plant construction and used a production cost model to calculate market prices. Interestingly, their calculations revealed a bust in market prices in both scenarios. They concluded by emphasizing the problem of covering the total levelized cost of a new CC with the average annual revenues earned in the energy market:

> *A new generator's profitability will depend largely on the prices it is paid for energy during the summer peak demand season, if it is relying solely on the energy market for revenue. Market clearing prices during the summer peak demand season may not reach a level necessary to sustain new market entry until reserve margins drop below historic levels usually regarded as necessary for reliable service.*

This conclusion is of fundamental importance, and it has been reemphasized in the recent <u>California Energy Outlook</u> (CEC 2001, Appendix B) and in the announcement for this workshop.

My simulation of power plant construction was expanded in the summer of 2000 with a model to represent construction and market prices in a summer peaking system with approximately the same loads and resources as those in California. The model simulated market prices on an hour-by-hour basis to represent the combination of the PX day-ahead price and the ISO real-time price for energy. The new model allowed for several types of investors who would apply for permits and commit to construction. The simulations revealed that construction could appear in a steady, even fashion, causing power plants to come on line exactly in time to meet the profitability goals of the investors. But this was not the dominant pattern. The more likely pattern showed construction lagging behind the growth in demand, allowing prices to climb to surprisingly high values during peak periods in the summer. When power plants are completed, they tend to come on line in great numbers causing a bust in wholesale prices (Ford 2001). The previous article concluded that the lack of power plant construction is a western problem (not just a California problem), and it called for an expansion in the model boundary to include loads and resources throughout the west. The expanded model, the "Western Market Model," is used in the remainder of this paper.

The Western Market Model

The western market model was constructed using the system dynamics approach which proved useful in the previous models of power plant construction. System dynamics was pioneered by Forrester (1961) and is explained in texts by Ford (2000) and Sterman (2001). It is valued as a strategic tool for a rapidly changing electric industry with high uncertainty and high risk (Dyner and Larsen 2001). The western market model is designed for highly interactive use to promote experimentation and discussion. The goal is general understanding, not precise forecasting.

Figure 2 shows the opening screen of the model with one of the information buttons in view. The button explains that the model operates as if the entire loads and resources in the WSCC interact in a single market place. The fundamental assumption is that market prices at various points in the west will rise and fall together (EPRI 20001). Support for the one market approach appears in the testimony by Oliver (2000), in the NPPC (2000) study of western markets and in FERC's June 19th Order.

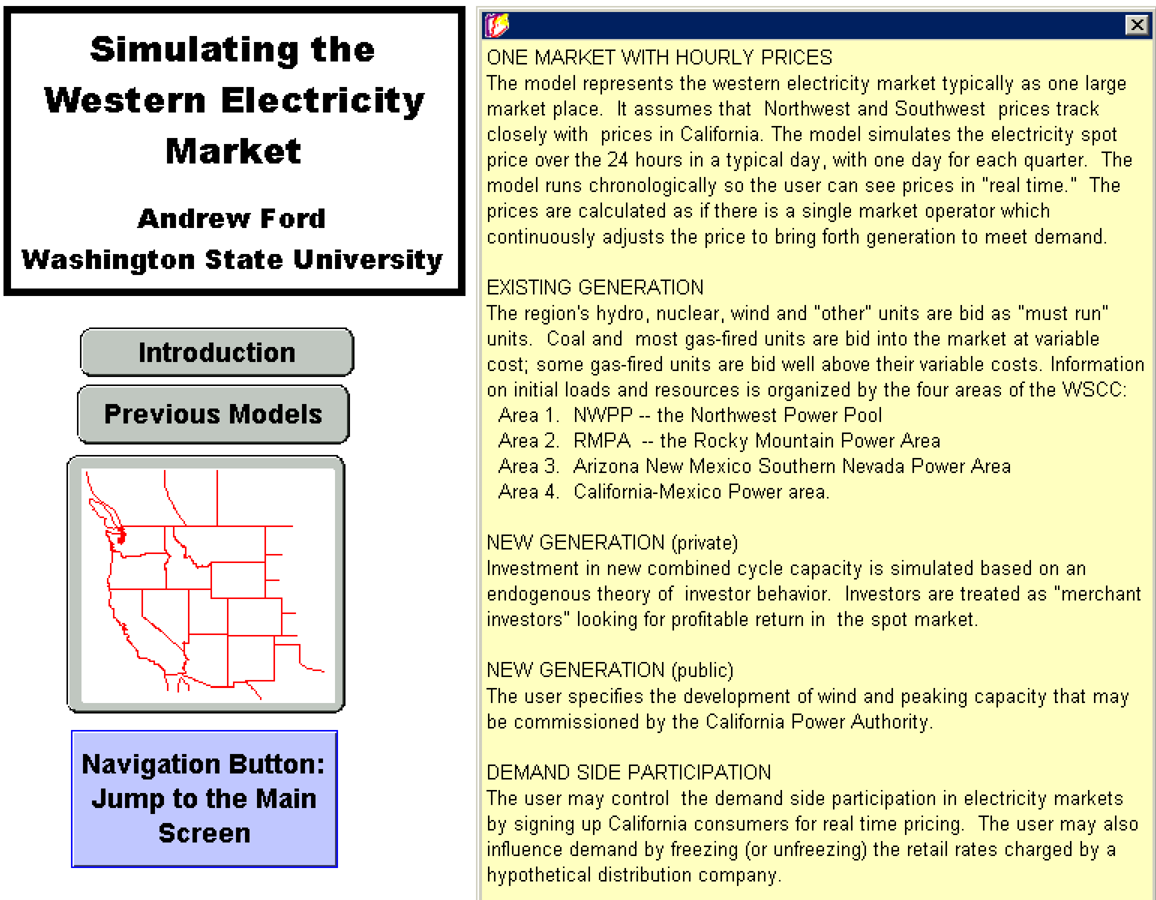


Figure 2. Opening Screen of the Western Market Model.

The simulations begin in the winter of 1998, with information on existing generation taken from the WSCC. The model simulates a typical 24 hour day for the winter of 1998, records the results as representative of the winter quarter, and proceeds to simulate a typical day for the spring of 1998. (Further information on the hourly operations is given in Appendix A.) The model simulates the interval from 1998 to 2001 before proceeding into the future. The historical period is simulated to allow one to compare simulated results with historical results. It's also useful to simulate the historical period to give the model time to develop the proper "momentum" as it enters the future.

Simulated Prices for the Past Four Years

Figure 3 shows simulated market prices over the interval from 1998 to 2001 with the vertical axis scaled from 0 to 400 $/mwh. The hourly prices tend to increase in the day and fall at night, but these hourly variations are not readily apparent in 1998 and 1999 because of the scale. The hourly variations are much more discernable during 2000 and 2001. The first price spike appears in the spring of 2000 (as highlighted by the button). The price peaks at nearly 200 $/mwh during the typical day in spring of 2000. Figure 3 shows larger and more persistent spikes in the remainder of 2000 and in the first half of 2001.

The model calculates averages over the 24 hours in a typical day for each quarter. These quarterly results appear in Figure 3 as abrupt changes when the model posts new results at the conclusion of each quarter. The quarterly price climbs to around 130 $/mwh in the summer of 2000 and even higher in the fall of 2000 and the winter of 2001. The winter price is around 250 $/mwh, nearly ten times higher than prices at the start of the simulation.

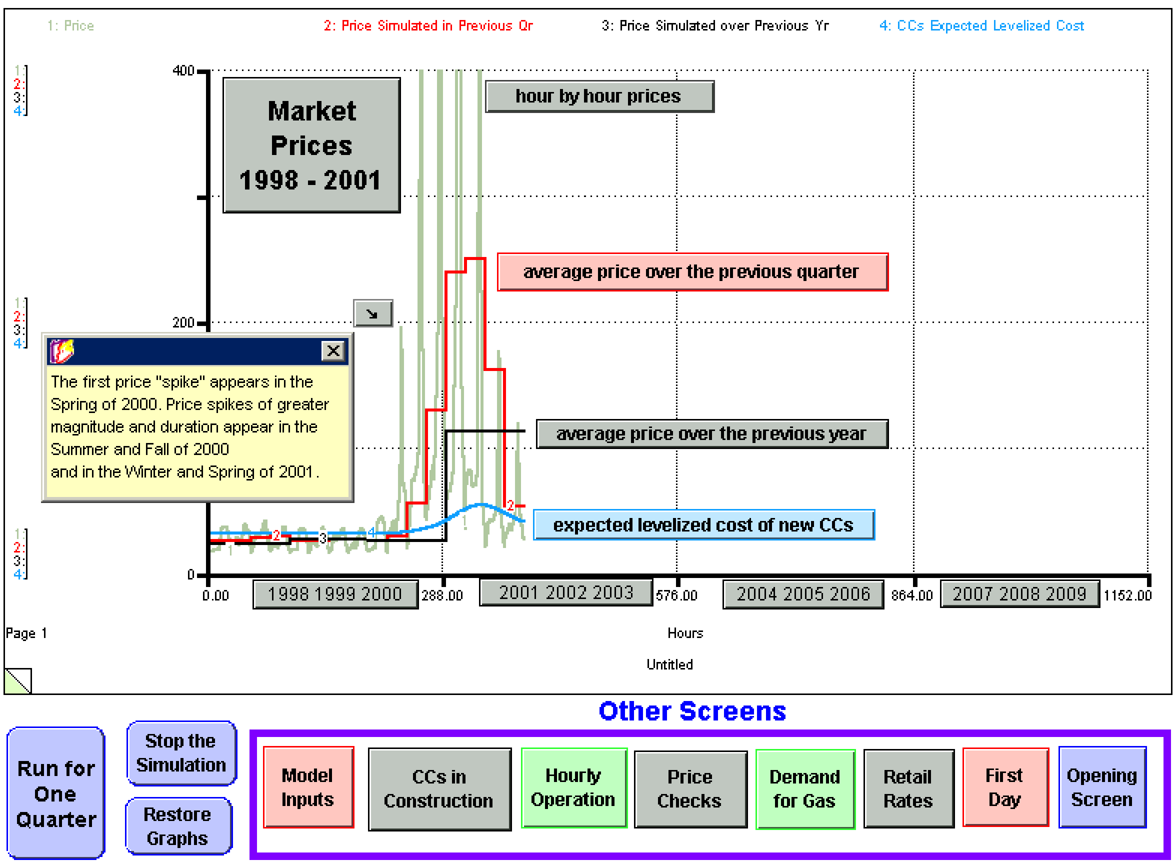


Figure 3. Main Screen with four years of Results.

These prices may seem shocking, but they are similar to quarterly prices reported by the ISO (see Appendix B for more information). The model calculates averages over the entire year as well. The annual results appear in Figure 3 as abrupt changes when the model posts new results at the conclusion of each year. Figure 3 shows an average of 27 $/mwh in 1998, 26 $/mwh in 1999. The average for 2000 is 113 $/mwh.

The final curve in Figure 3 is the investors' expected levelized cost of a new CC. This is displayed along side of the market prices to provide perspective. In the first two years, the model shows investors' expectation for a new CC at 31.5 $/mwh. This is significantly higher than the average annual market clearing price in the first two years. The price of natural gas increases dramatically in 2000, and the investors react with an upward adjustment in the expected levelized cost of a new CC. Figure 3 shows their expectation peaking at around 54 $/mwh mid way through 2001. This value turns out to be less than half the average annual price for 2000, and it is nearly five times smaller than the average price in the winter of 2001.

The lower portion of the main screen is filled with operational and navigation buttons. There are buttons to run or stop the simulation and a row of buttons to jump to other screens for setting the inputs or for viewing the results. Figure 4 shows the main screen for setting model inputs.

Initial loads are based on the WSCC data for each reporting area, and the user specifies the growth in demand using the four sliders in the upper left portion of Figure 4. The historical simulation assumes a general trend of 2 % annual growth in each of the four years. Variations from this trend are imposed to account for changes in the weather. For example, the summer of 1998 was somewhat hotter than normal; the summer of 1999 was somewhat cooler than normal.

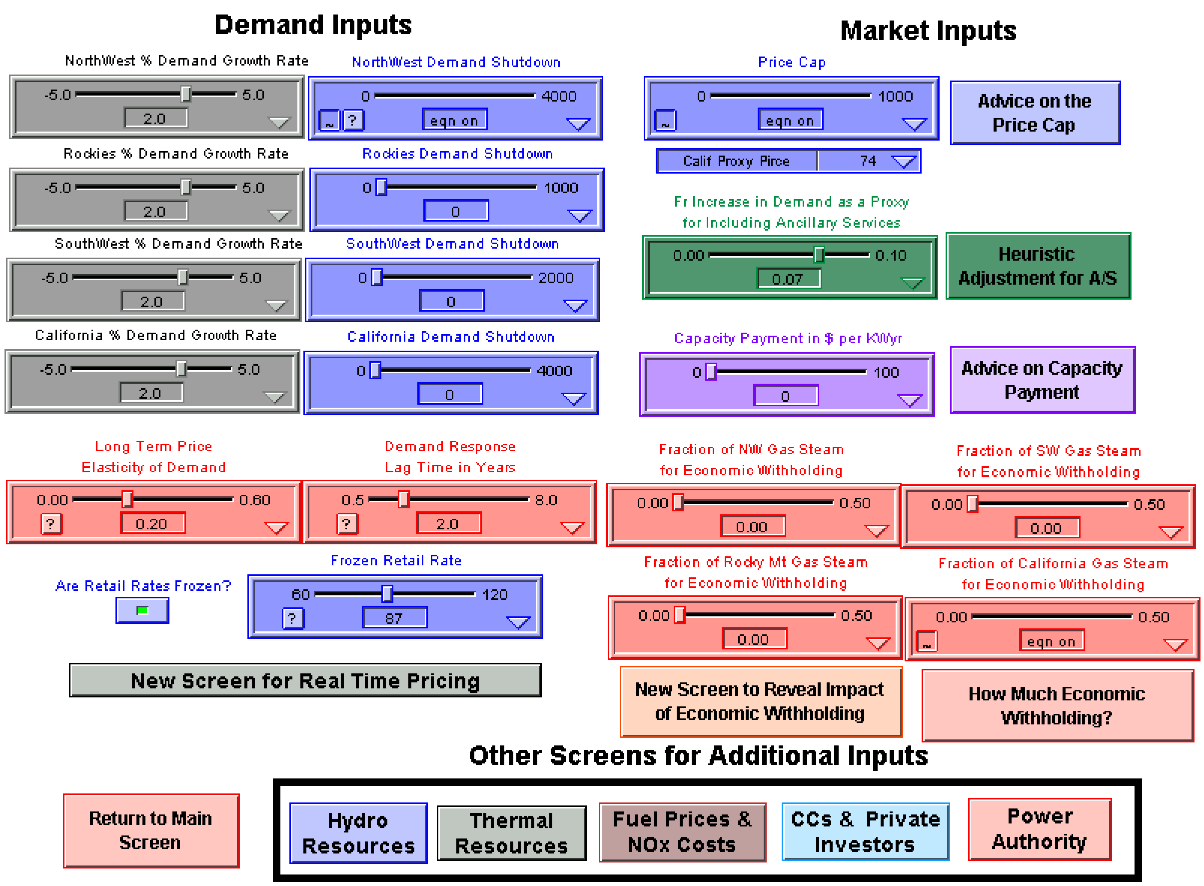


Figure 4. Model Inputs Screen.

Figure 4 shows four additional controls to allow one to impose the unusual reduction in loads that have appeared during the past few years. The top control is used in this simulation to lower northwest loads near the end of 2000 and into 2001. This represents the major shutdown in industrial loads (such as the aluminum smelters normally served by Bonneville).

Additional “demand inputs” in Figure 4 are the price elasticity of demand and lag for consumers to respond to changes in retail prices. The long-term price elasticity is set at 0.2 (which means that consumers would lower their demand by 20% if retail prices were to increase by 100%). The retail rate for the hypothetical distribution company is frozen at 87 mills/kwh, however, so there is no consumer response in this simulation.

The right side Figure 4 shows the key parameters for representing market design and market behavior. The first input is the price cap, expressed in $/mwh. I assume that the cap in the ISO real-time energy market serves as the defacto price cap in the California markets and for the entire western market. The historical simulation begins with the cap at 250 $/mwh. The cap was raised to 750 $/mwh in October of 1999, lowered to 500 $/mwh in July of 2000, lowered again to 250 $/mwh in August of 2000. These and other changes are represented by the "equation on" control for the price cap. A Proxy Price (similar to FERC's use of a proxy price for "price mitigation") is shown for comparison purposes.

The next control in the market inputs column is the "fractional increase in demand used as a proxy for including ancillary services." The base value is 7%, which means that the actual demand is elevated by 7% before calculating market prices. The generating resources are bid into a market to serve the elevated demand, and the price is taken as an approximation for the energy price that would result when generators can bid into multiple markets. The slider allows this input to vary from 0 to 0.10. The value of 0.10 is appropriate if one wishes to be "conservative" in estimating market prices. The conservative approach was followed in FERC testimony by Kahn (2000) and in calculations by Hildebrandt (2000). Their conservatism was appropriate because they were estimating the extent of market power. For our purposes, however, it makes more sense to set the adjustment at a more realistic factor. I selected a 7% adjustment which corresponds to the WSCC guideline on reserves.

The remaining sliders shown in Figure 4 allow the user to control the extent of "strategic behavior." Strategic behavior usually takes the form of physical withholding or economic withholding. The Western Market Model is designed as if the impact of strategic behavior can be represented by a user specified fraction of older gas units which are subject to economic withholding. To simulate competitive outcomes, one sets these fractions to zero, and all of the gas capacity will be bid at variable cost. Turning the California "equation on" assumes around 10% withholding during 1998, higher fractions in 2000. These "default values" were selected to give realistic prices, as explained in Appendix B.

Simulated Construction for 1998-2001

Figure 5 shows the results on the "CCs Under Construction" screen. The four variables are displayed on a graph scaled from 0 to 60,000 MW. The "paperwork on proposed CCs" grows to almost 45,000 MW by the end of the historical period. The first gray button is located to represent the 33,000 of proposed projects (either approved or in the formal review process) at the end of 2000. The next gray button shows a historical benchmark of 43,000 for paper work midway through 2001. The simulated accumulation of paperwork comes close to the two benchmarks.

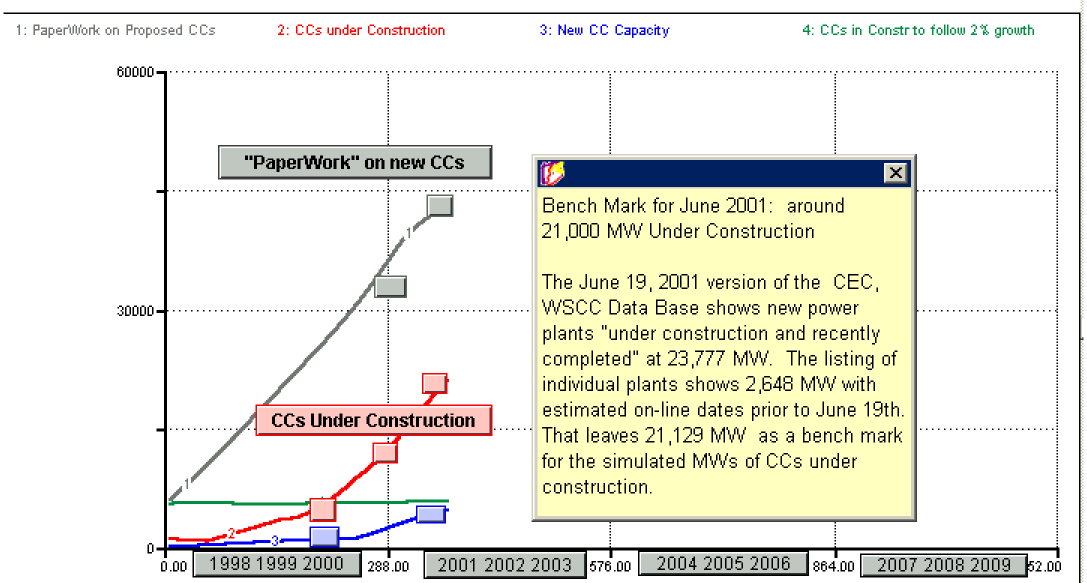


Figure 5. Simulated Construction Over the Past 4 years.

The three red buttons represent benchmarks for new CCs under construction midway in 2000, at the end of 2000, and midway through 2001. The third button is open, so we can read that around 21,000 MW of capacity was under construction midway through 2001. Figure 5 shows that the simulated growth in CCs under construction comes close to all three benchmarks.

The two blue buttons show the total new capacity that has completed construction during the historical period. The first button represents around 1,000 MW on line by the start of 2000; the second button represents around 4,000 MW by summer of 2001. Figure 5 shows that the simulated growth in new capacity comes close to the two benchmarks.

The fourth variable displayed in Figure 5 shows the MW of CCs that would be in construction if investors were building new power plants to expand total generating capacity to keep pace with the 2% annual growth in demand. The construction interval for a new CC is around 24 months (EPRI 2000). Given this lag and given the size of the WSCC system, we would need to see around 5,500 MW under construction. Figure 5 shows that the simulated construction is well below the 5,500 MW in 1998 and 1999, well above in 2000 and 2001. In other words, we can summarize the historical results by saying that investors were

- Under-building in 1998-1999 and
- Over-building in 2000-2001.

The under-building in 1998-1999 is one of the main factors contributing to the severe price spikes in 2000-2001. One might wonder why we have not seen more MW of construction completed in time for their owners to benefit from the price spikes in the summer of 2000. With a 24 month construction delay, investors would have had to start construction midway through 1998 if they were to bring their plants on line midway through 2000. Perhaps they did not start construction at this time because of “opening market confusion.” After all, the California markets did not begin operation until the spring of 1998. On the other hand, one might argue that market rules were evident with the passage of AB 1890 in the summer of 1996. With the rules in place, a rational investor might have looked into the future and forseen the large profits to be earned from starting construction midway through 1998.

The explanation of the historical under-building will shape the way one addresses the question of changing California’s wholesale market structure for the future. If one attributes the under-building to to “early market confusion,” one might argue that we should retain the current market structure and hope the investors will be less confused in the future. On the other hand, if the under-building is attributed to a combination of factors that could reappear in the future, then we need to examine new market structures. The Western Market Model explains the under-building based on market fundamentals that could well reappear in the future.

A Theory of Investor Behavior

The diagram in Figure 6 shows the theory of investor behavior implemented in the Western Market Model. The model simulates the development process beginning with the application for a construction permit. After 12 months, the developer receives the permit, and the project enters a “site bank”. (The combination of CCs under review and in the site bank was shown in Figure 5 as “paperwork on proposed CCs.”) Approved projects give the developer an option to start construction, but there is no requirement that the developer must follow through on the proposed project.

The key decision is whether to start construction, and Figure 6 depicts the model’s theory of construction starts with some illustrative numbers that would apply during a period like 1998-1999 when natural gas was priced at around 2.50 $/mmBTU. Figure 6 shows the three assessments that investors conduct to arrive at the simulated construction starts:

- **Supply Assessment:** The model assumes that investors are aware of all of the capacity in the WSCC. They know the planned and scheduled outages of the thermal units, so they are able to estimate the thermal capacity during a peak day in the summer. The investors count the new CCs that have come on line in their assessment, but they may not count all of the CCs that are currently under construction.

- **Demand Assessment:** The model assumes that investors watch the trend in demand over time, and they use the trend to estimate peak demands in the future. Their forecast looks two years in the future since it takes two years to build a new CC. The forecasted peak demand is compared to the expected supply to get the expected reserve margin. In this example, investors are expecting a future reserve margin of 15%.

- **Market Assessment:** The model assumes that investors have access to production costing models that would allow them to calculate average annual market prices based the reserve margin. With a 15% reserve margin, for example, they might expect the future market to clear at 26 $/mwh. The model uses a highly nonlinear curve to represent the changes in estimated market prices based on the expected reserve margin. This particular curve is based on market price calculations with natural gas priced at 2.50 $/mmBTU.

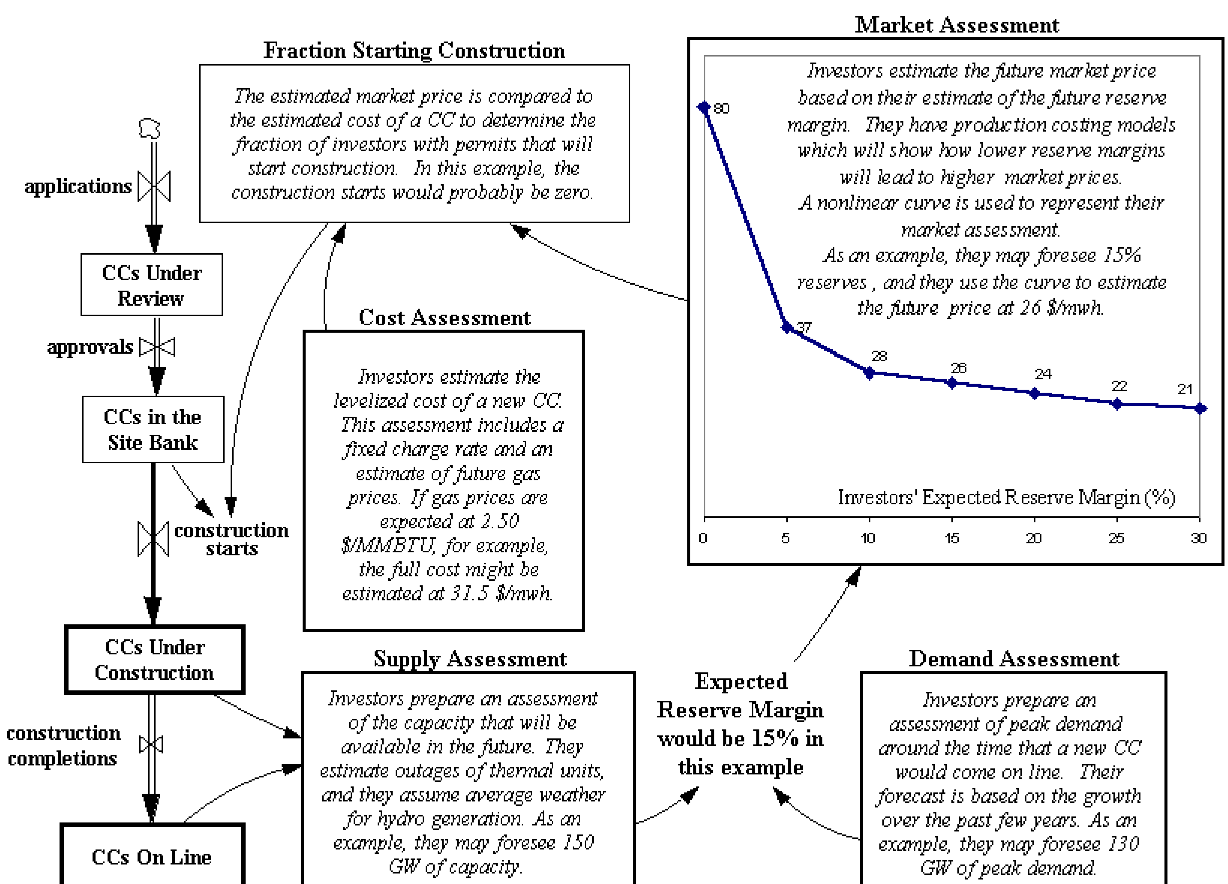


Figure 6. The Theory of Investor Behavior Implemented in the Western Market Model.

In the illustrative example, the investors expect the market to clear at 26 $/mwh, but the full cost of a new CC is expected to be around 31.5 $/mwh. The model assumes some diversity of investor conditions, so some investors may start construction when expected market prices are somewhat below 31.5 $/mwh. But in this example, 26 $/mwh is simply too low for a significant fraction of the developers to begin construction. The model assumes that investors would be inclined to wait for expected conditions to improve. With time, demand will grow, expected reserve margins will fall, and expected market prices will rise. When expected market prices are closer to the investors' target for a new CC, they will turn their permits into actual construction projects.

The approach in Figure 6 is simulated continuously over time. That is, investors continuously update their assessments of supply and demand as simulated conditions change over time. If they do start construction, their own construction will shape their assessments in the future. This approach succeeds in explaining construction over the historical period. It explains the under-building in 1998-1999, and it does so without resorting to the argument that investors were inhibited by "early market confusion." It also succeeds in explaining the over-building in 2000-2001. The next section shows the implications of extending this simulation into the future.

The Business As Usual Scenario

Several of the model inputs were adjusted from year to year during the historical period. For example, the price of natural gas increased rapidly in 2000 and early 2001 before declining midway through 2001. As another example, hydro generation was higher than average in 1998 and 1999 but well below average in 2000 and 2001. These variations are helpful during the historical period because they improve the comparison with historical prices, as explained in Appendix B.

However, as we extend the simulations into the future, it makes sense to adopt relatively constant assumptions in the interest of clarity. For example, I assume that hydro generation will be based on "average" conditions for the future. I assume natural gas will cost 4 $/mmBTU in California and 3 $/mmBTU in the rest of the WSCC. The thermal units will experience fixed outage rates, and the planned maintenance will be scheduled during the same seasons as in the past. Although there are many older power plants in the west (especially in California), there are no retirements of existing capacity in the business as usual scenario.

New Capacity	**Demand**
CC Capital Cost = 600 $/kw	Initial Demands = WSCC data
CC Fixed Charge Rate = 14.5 %/yr	Demand Growth = 2 %/yr in all 4 areas
CCs Fixed O&M = 10 $/yr per kw	Demand Shutdowns = none
CC Heat Rate = 6,800 BTU/kwh	Customers on Real-time Pricing
CC Permitting Interval = 12 months	Programs = none
CC Permit Shelf Life = 10 yrs	Retail Rates = frozen
CC Developers' Goal	
for Paperwork = 45 GW	**Market Operations**
CC Developers' Diversity = 6%	Adjustment in Demand for A/S = 7%
CC Construction Interval = 24 months	Capacity Payment = none
CC Developers' Attention to the	Economic Withholding =
Construction Pipeline = 50%	20% of older gas capacity in CA
New Peakers (CTs) = none	QFs Shutdown (credit crisis) = none
New Wind = 530 MW in NW	Price Cap = 200 $/mwh
Existing Capacity	**Fuels and Emissions**
Thermal Capacity = WSCC data	Natural Gas Price in California
Forced Outages = fixed	= 4.00 $/mmBTU
Scheduled Outages = fixed	Natural Gas Price in NW, RM, SW
Hydro Capacity = WSCC data	= 3.00 $/mmBTU
Hydro Generation = "average year"	Cost of NOx Credits = 0 $/mwh
Hydro Shaping = fixed	Coal Price in RM, SW = 0.75 $/mmBTU
Retirements = none	Coal Price in NW = 1.00 $/mmBTU
	Coal Price in CA = 1.25 $/mmBTU

Table 1. Long-Term Assumptions for the Business As Usual Scenario.

The new capacity will come from private investment in the CCs. There are no investments by the California Power Authority, but I do include Bonneville's commitment to 530 MW of wind capacity in the northwest. The scenario assumes that a price cap will remain in place and that economic withholding will remain at the values found useful in explaining historical prices. There are no capacity payments; no real-time pricing programs; and retail rates are frozen

Simulated Construction

Figure 7 shows the simulated construction in the business as usual scenario. The scenario envisions that investors will wish to maintain the paperwork at the amount accumulated over the past few years. The total paperwork remains at approximately 45,000 MW from 2002 through 2009. This is a combination of projects under review as well as projects in the site bank. As permits are granted, the bulk of the paperwork will be in the site bank.

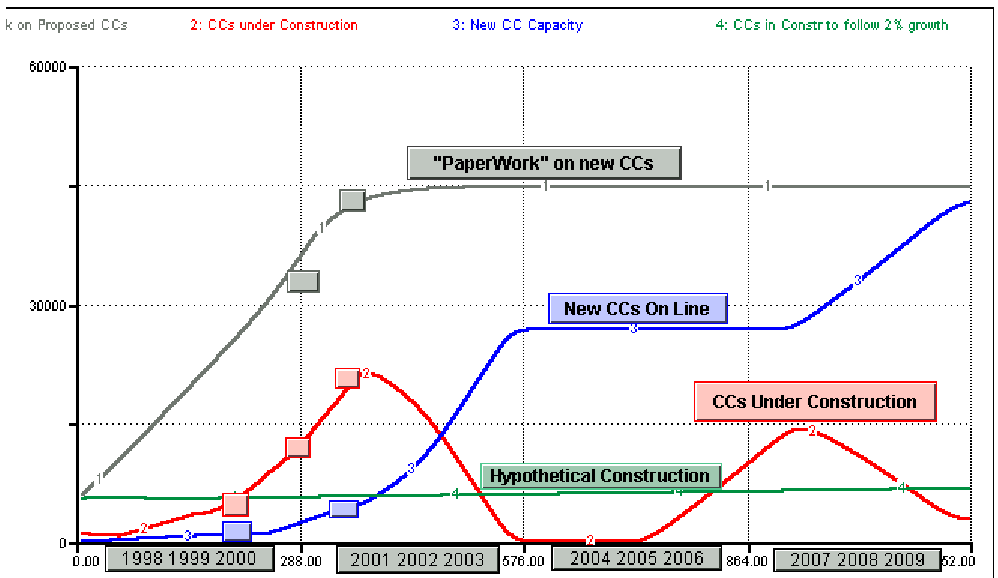


Figure 7. Simulated Construction in the Business As Usual Scenario.

The red curve in Figure 7 shows that CCs under construction peaks near the end of 2001. From this point forward, construction completions are greater than construction starts, and the total MW of CCs under construction declines. This simulation suggests that we are now at the crest of the current building boom. As the construction is completed during 2002 and 2003, installed CC capacity will grow to around 27,000 MW by the start of 2004. Installed capacity remains at this value for the next three years because of a lull in construction during 2003, 2004 and 2005. Figure 7 shows a second wave of construction beginning in 2006 and peaking near the end of 2007.

The green curve shows the "CCs in Construction to follow 2% growth." This is the hypothetical amount of construction that may be used to judge whether investors are over-building or under-building. We are currently in the midst of a building boom, and the actual construction is well above the hypothetical construction. This over-building situation continues until midway through 2003. For the next three years, however, the scenario envisions construction below the levels needed to keep pace with demand. This lull in construction arises when investors again arrive at pessimistic market assessments (similar to the assessment depicted in Figure 6). During this interval, investors are reluctant to start construction, even though they hold a huge number of approved permits. According to this scenario, investors hold off on construction starts until 2006 – 2007. This second wave of construction causes installed CC capacity to grow again in 2008-2009. By the end of the simulation, there is 43,000 MW of installed CC capacity in the west.

Simulated Prices

Figure 8 shows market prices in the business as usual scenario, with the vertical scale running from 0 to 400 $/mwh as in the previous display. The simulation indicates that hourly variations for typical days in 2002 and 2003 would be much smaller than the variations in 2001. Figure 8 shows the quarterly prices continue to decline during the interval from 2002 to 2004.

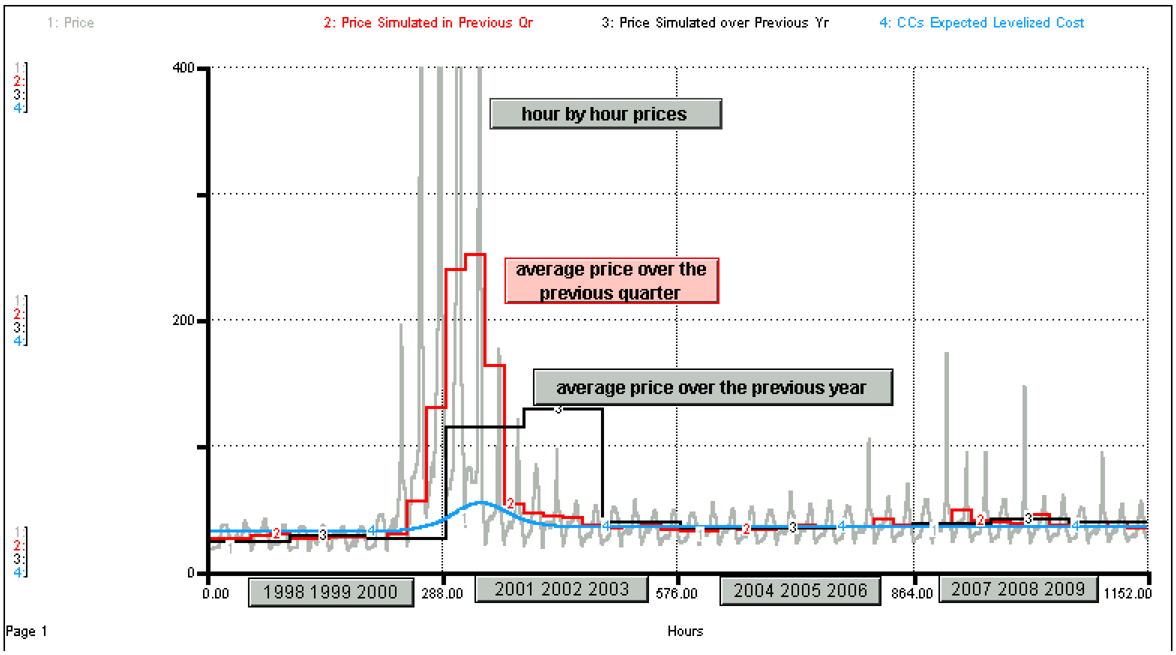


Figure 8. Simulated Prices in the Business as Usual Scenario

Figure 8 shows the average annual price in black. The annual price is updated at the end of each year, and the results appear as abrupt changes in Figure 8. The average price for 2000 is 114 $/mwh; the average for 2001 is even higher, at 129 $/mwh. The construction boom allows for much lower prices for the interval from 2002 to 2005. The average in 2002, for example, is 39 $/mwh. By 2003, the average annual price is down to 34 $/mwh. At this level, the annual price is slightly below the investors' expected cost of a new CC (which is at 35 $/mwh).

An important result from the business as usual simulation is the reappearance of price spikes in the years 2006 and 2007. The spike in 2006 exceeds 100 $/mwh; the spike in 2007 is around 175 $/mwh. These contribute to an increase in the quarterly and annual prices. In 2007, for example, the average annual price is 41 $/mwh, somewhat above the investors' expected cost of 35 $/mwh for a new CC. The magnitude of these price spikes is much lower than the spikes during 2000 and 2001, but their timing is similar. The spikes in 2000 and 2001 appeared in the midst of a large building boom, just prior to the majority of the new CCs completing construction. The spikes in 2006 and 2007 appear in the midst of the second building boom, just prior to the next major increase in installed capacity.

Simulated Reserves

Figure 9 shows the simulated reserves in the business as usual scenario displayed on a scale from – 40% to 120%. The Power Authority has declared the need to maintain reserves at 15% or higher, and this target is displayed for comparison. The reserves that would trigger a Stage I alert, a Stage II alert or a Stage III alert are also shown for comparison. The simulated "reserve margin" varies dramatically during the course of a day. This variable is defined as the current operating reserves divided by the current demand. Operating reserves are defined as all of the generation the system operator may call upon during each hour of the day.

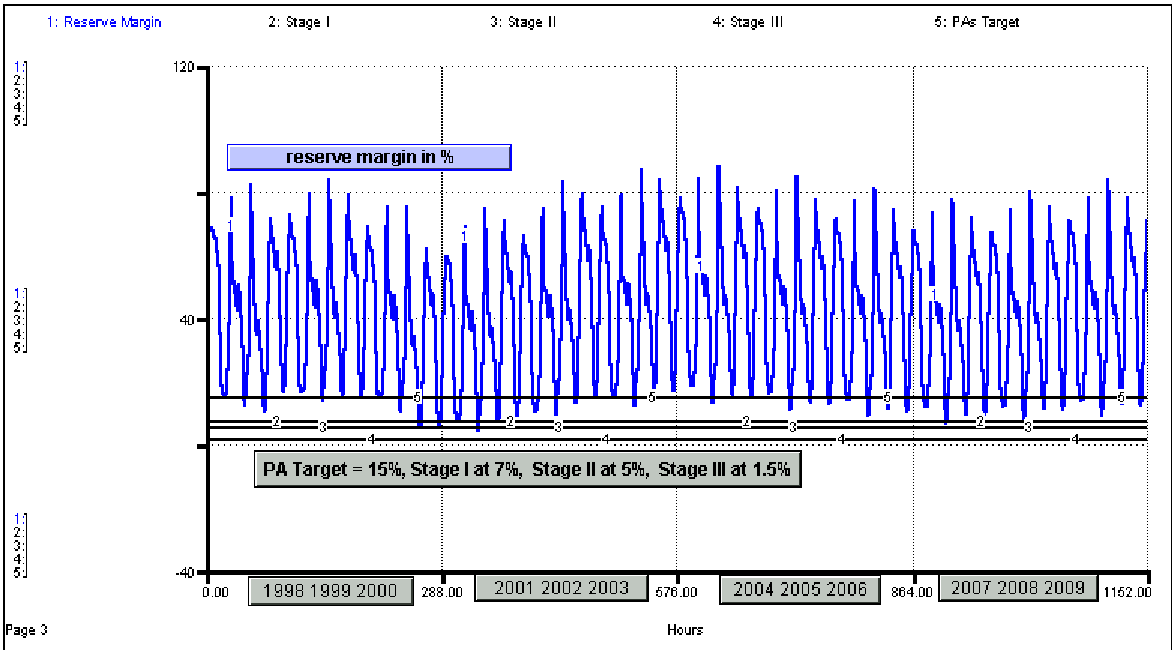


Figure 9. Reserves in the Business as Usual Scenario.

Figure 9 shows the reserve margin falling to its lowest value when demand climbs to a peak value each day. For our purposes, the important comparison is the minimum reserve margin in each day compared to the 15% target by the Power Authority or compared to the values which would trigger an alert. Figure 9 shows that the Power Authority goal is seldom achieved, either in the past or in the future. Figure 9 shows that alerts would be declared in the year 2000 and 2001, but the building boom allows reserve margins to climb above 7% during the interval from 2002 to 2006. By 2007, however, reserve margins have returned to the low levels that would force the system operator to declare alerts.

Commentary

An important conclusion from Figure 9 is that the western system will not provide the reserve margins of 15% said to be necessary for a reliable system. Reserves would dip below 15% even during 2003 and 2004, the years immediately after the completion of the first wave of construction. By 2007 and 2008, reserve margins are back at unreliable levels, reminiscent of the situation in 2000 and 2001.

A second conclusion is that construction would appear in repeated waves of boom and bust. Figure 7 shows a second wave of construction cresting in the year 2007, around six years after the crest in the first wave of construction. The pattern in Figure 7 is quite regular, but this regularity comes from the constant assumptions adopted in this scenario. With more a complicated and erratic scenario, the waves of construction would appear in an irregular fashion.

A third conclusion, is that price spikes could reappear as soon as 2006 and 2007. These spikes are much less severe than the spikes in 2000 and 2001. The improved behavior arises because the scenario does not envision a replay of the "perfect storm." For example, gas prices do not skyrocket in 2006 like they did in 2000. Also, hydro generation does not decline in the years 2006 and 2007, like it did in 2000 and 2001. These assumptions mask the underlying deterioration of the supply-demand balance that occurs during the lull in construction.

Figure 10 shows a variation in the opening scenario to unmask the difficult conditions that could emerge within the next decade. The new scenario makes two changes in the previous assumptions. First, I assume that the northwest could experience a “dry year” in 2007 (the region would lose around 6,000 average MW of generation). The dry year is timed to appear at approximately the same point in the boom-bust cycle as the dry conditions in 2000 and 2001. The second change is to lift the price cap at the start of 2004. We might envision the cap would be lifted because those who oppose price caps would win adherents to their view during the previous two years, years of low prices and declining construction.

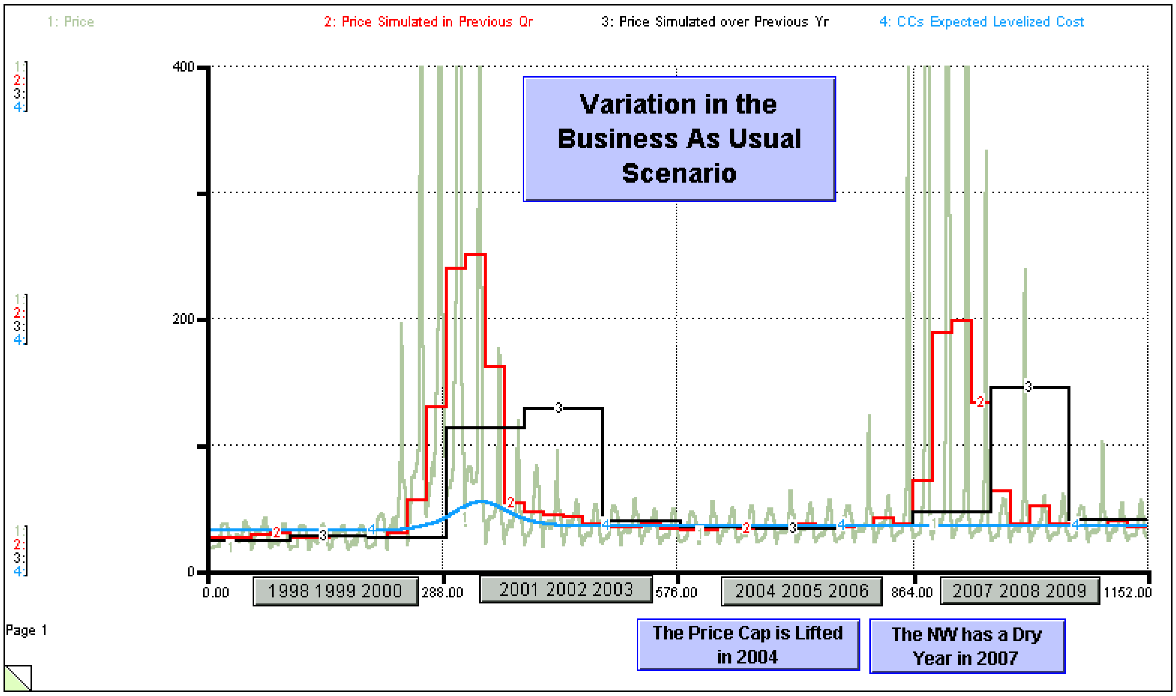


Figure 10. Market Prices in a Business As Usual Scenario with the Price Cap Lifted in 2004 and a Dry Year in the Northwest Hydro System in 2007.

Figure 10 shows hourly prices spiking “off the chart” in the year 2007. Quarterly prices would climb to around 200 $/mwh, and the average annual price for 2007 would jump to 146 $/mwh, a value which exceeds the annual prices seen in either 2000 or 2001. Reserve margins (not shown here) would decline to levels lower than the simulated reserves in the years 2000 and 2001.

The price spikes in Figure 10 dramatize the problems inherent in the boom and bust pattern of construction. The simulation reveals that the western system could be “one dry year away” from a repeat of the crisis conditions that appeared in the years 2000 and 2001. If we do not implement fundamental changes in the structure of wholesale markets, we run the risk of exposing the west to another round of price spikes and rolling blackouts.

Supply-Side Scenario #1:
New Incentives for Timely Construction

One of the factors contributing to the poor behavior in Figure 10 is the lack of timely construction by private developers. The price spikes in 2000 and 2001, for example, may be attributed to the under-building during 1998 and 1999. The spikes in 2006 and 2007 may be attributed to the under-building in 2004 and 2005. The fundamental factors that lead to the under-building have been depicted in Figure 6. During years when investors foresee adequate reserve margins they would expect market prices to clear below the fully levelized cost of a new CC. Under these conditions, they hold back on construction, waiting for market conditions to improve. By holding back, they set the stage for a larger building boom, one which will come too late to ensure that the west has sufficient generating capacity. This system needs new incentives to encourage investors to start construction earlier.

Background on Capacity Payments

Capacity payments provide incentives for generators to be available when the system needs capacity, and they provide "extra revenue to the generator to cover the capital and other fixed costs which are not covered by the energy price" (Hunt 1966, p 111). Many have argued that capacity payments are needed in the west (CERA 2001, McCullough 1998, Michaels 1997), and I have advocated their adoption in a previous modeling study (Ford 1999). On the other hand, several studies are silent on the topic of capacity payments (CBO 2001, CATO 2001, ERPI 2001). At this point, the California markets allow capacity payments only in the form of short-term payments for ancillary services.

Some may be opposed to the introduction of capacity payments because of the unfortunate experience in the early years after the UK decision to privatize their electric system. Their original pool established a capacity payment at approximately the product of the loss of load probability and the value of the loss of load (set at 3,000 $/mwh). This payment was important, contributing around 20% of the generators' earnings in 1994 and 1995 (Newbery 1995). Unfortunately, it was extremely sensitive to reserve margin, making it vulnerable to manipulation by owners with horizontal market power, and it was eliminated in 1998 as part of a general overhaul in their trading arrangements. For our purposes, the work by Bunn and Larsen (1992) on the UK payment is instructive. They used computer simulation to demonstrate that a variable capacity payment would contribute to highly volatile patterns of construction. The simulated volatility did not arise from the exercise of market power. Rather, it was due to combination of factors including the size of the payment, the cost and lead-time of new CCs and the limits on investors' ability to see into the future.

It's evident from the business as usual simulation that the last thing we need in the western markets is additional volatility. This system needs additional incentives, not more volatility. We should avoid a variable capacity payment that could rise or fall with swings in the reserve margin. In the interest of stability, it makes better sense to provide the incentive in the form of a capacity payment that remains constant over time. The recent California Energy Outlook (CEC 2001) discusses the distinctions between energy and capacity. It then explains that capacity payments are one of the methods to "solidify the relationship between generation and load," and it shows the energy prices required to meet an investor's revenue requirements with capacity payments set at 20 $/kw per year or at 40 $/kw per year.

Simulated Impacts

The capacity payment control is shown in Figure 4. It allows the user to vary the payment from 0 to 100 $/kw per year. The high value would cover the entire fixed costs of a new CC (CEC 2001, p B-9). As an illustrative example, I have selected 20 $/kw per year, the smaller of the two examples studied by the CEC. I assume that investors use an expected capacity factor of 80% to include a $/mwh benefit in their profitability assessment. With this assumption, the capacity payments adds just under 3 $/mwh to the energy price. This may seem like a small incentive, but it could turn out to provide the extra boost that investors need to start building earlier. (To envision the impact, imagine a 3 $/mwh upward shift in the market assessment curve shown previously in Figure 6.)

Figure 12 shows the simulated construction if the capacity payment is implemented at the start of 2003, one year after the crest in the construction boom. Figure 12 indicates that the total CCs under construction would continue on a downward trajectory during 2003 despite the extra incentive. But the downward trend is reversed near the end of 2003. New construction starts exceed completions around this time, and the MW of construction climbs during 2004. The new scenario then shows total CCs under construction declining in a gradual fashion, but never falling below the "hypothetical construction" which would appear if construction were to expand capacity to keep pace with 2% annual growth in demand. The important conclusion from Figure 12 is that capacity payments could lead investors to start building earlier. This, in turn, leads to less over-building at a later point in the simulation. The overall impact is a substantial reduction in the tendency toward boom and bust in construction.

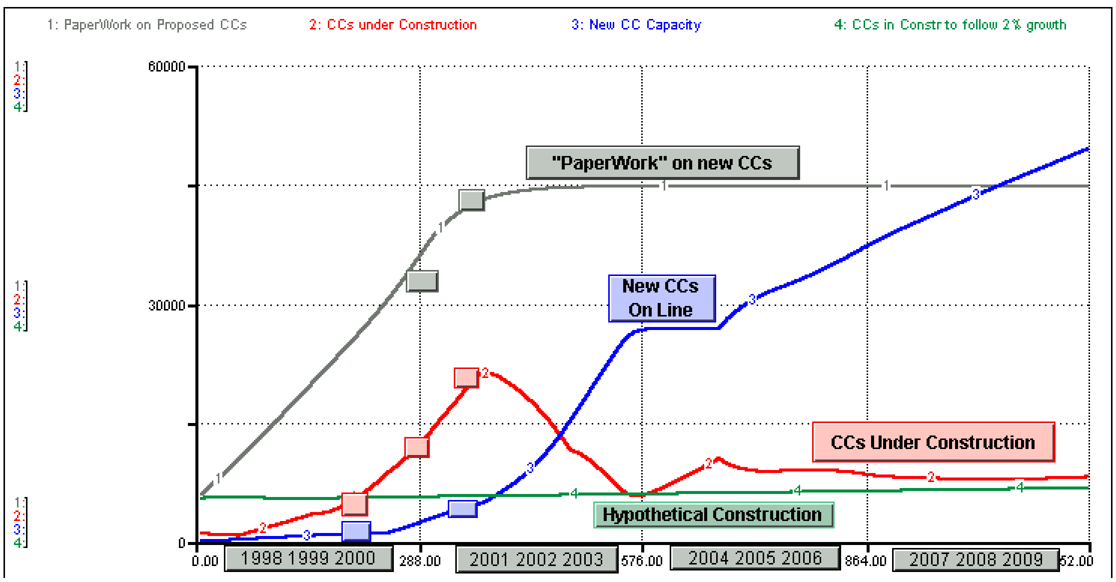


Figure 12. Construction in Supply Side Scenario #1:
Capacity Payments of $20 per kw-yr starting in 2003.

Figure 13 shows the reserve margins in the new scenario. As before, the best way to interpret this graph is to compare the minimum reserve in each day with the 15% target announced by the California Power Authority. The reserves in Figure 13 are somewhat improved (relative to the previous scenario) after the year 2005. By this point, the extra construction of new CCs allows the minimium reserves to dip only slightly below the target of 15%.

Figure 14 shows the market price implications of the capacity payments scenario. The "price" shown in gray in this display is the energy price. Figure 14 reveals that the variations in the energy price during the course of a simulated day would be greatly reduced by the introduction of capacity payments. The quarterly average prices and the annual average prices in Figure 14 are calculated from a combination of the energy price and the capacity payment. For purposes of comparison, the capacity payment is spread across 80% of the hours in the year, leading to an equivalent energy charge of just under 3 $/mwh. This is added to the market clearing energy price to get a total price for each hour. The average values of the total price are displayed in Figure 14, but the large scale makes it difficult to see the differences between this scenario and the previous scenario.

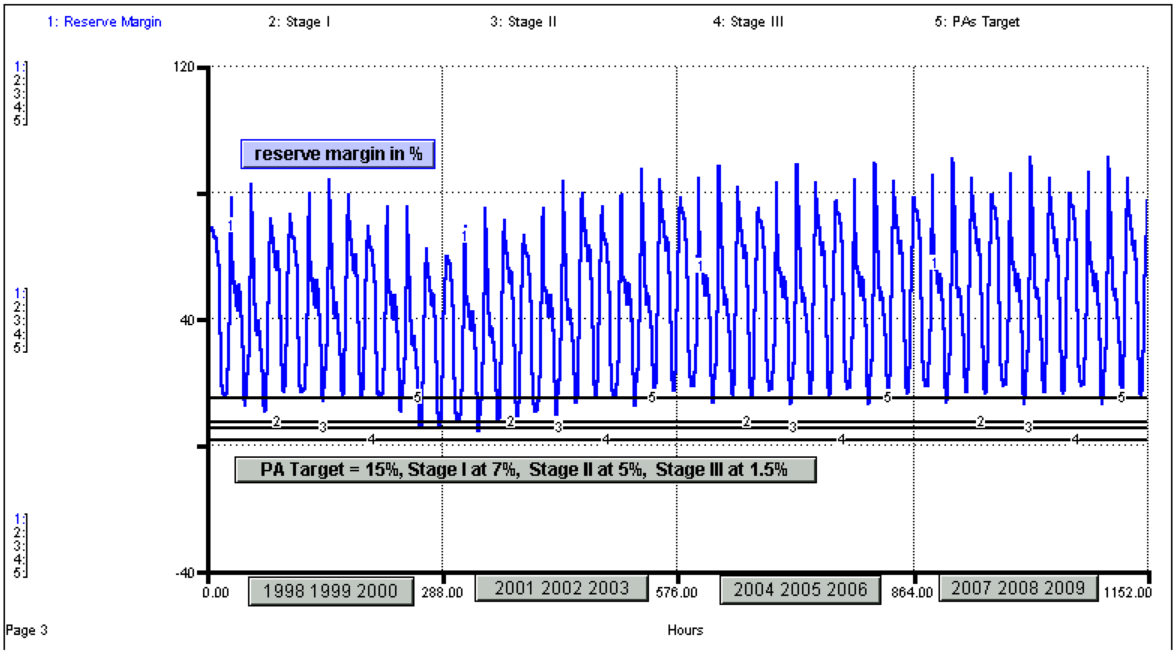


Figure 13. Reserves in Supply Side Scenario #1:
Capacity Payments of $20 per kw-yr starting in 2003.

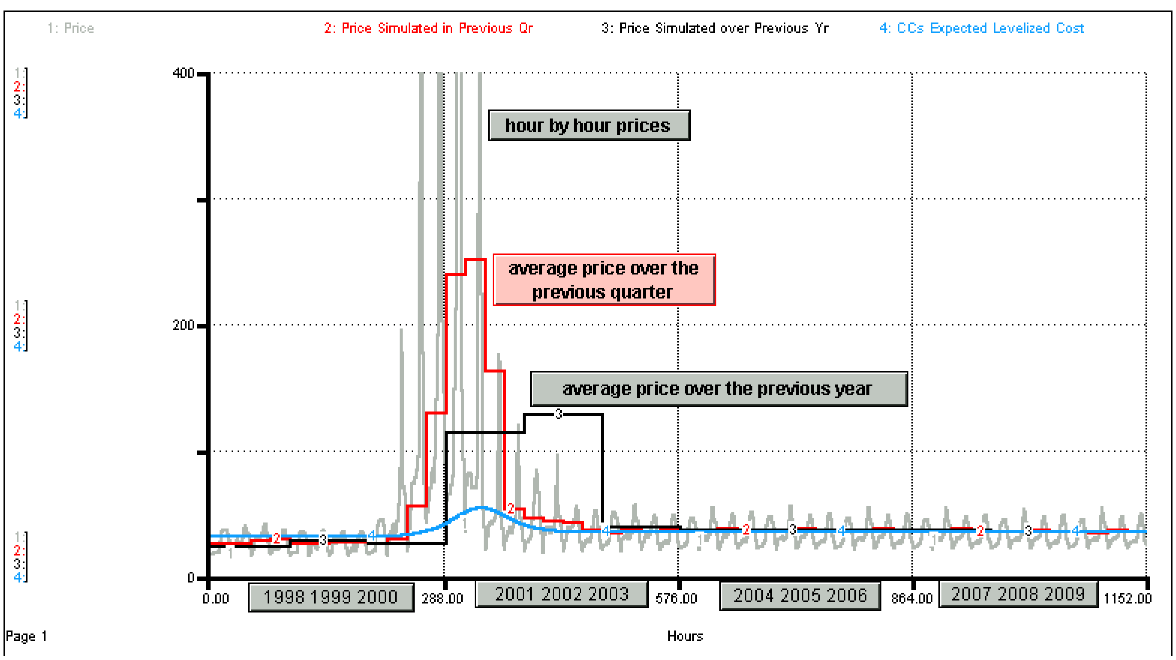


Figure 14. Prices in Supply Side Scenario #1:
Capacity Payments of $20 per kw-yr starting in 2003.

Figure 15 shows a comparative graph of the quarterly average prices from the two scenarios. The vertical scale has been reduced to 120 $/mwh to make the differences easier to see. Two buttons are added to aid in the interpretation. The higher prices button stretches from 2003 to 2006 to alert us that total prices would be higher during the first three years after the capacity payments are introduced. This is the time period when no additional generating capacity has come on line, so the market price of energy is the same as the previous scenario. The additional 3 $/mwh (which is paid to all generators) causes the increase in the total wholesale price.

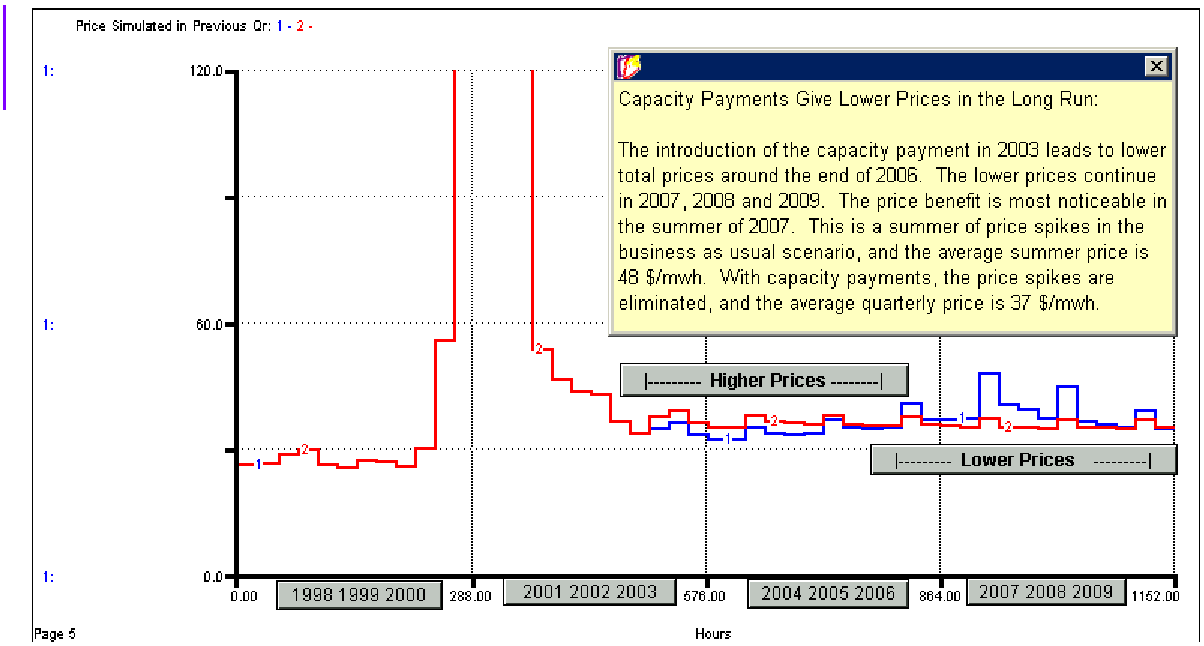


Figure 15. Comparison of Total Wholesale Prices:
Capacity Payments Lead to Higher Prices in the Short Run, Lower Prices in the Long Run.

The lower prices button is open in Figure 15. It stretches from 2006 until the end of the simulation. The reduction in prices is made possible by the extra generating capacity that is constructed in the new scenario. The additional capacity shifts the balance of supply and demand to such an extent that the system operator is able to meet the energy needs with prices that are more than 3 $/mwh below the energy prices in the previous scenario. The total price benefit is particularly evident in the summer of 2007, a time when price spikes appeared in the business as usual scenario.

Commentary

These results indicate that the problems of boom and bust could be reduced if private investors receive an additional incentive in the form of a fixed capacity payment. A modest payment of 20 $/kw per year is shown to break the cycle of boom and bust that would appear in the business as usual scenario and to lead to higher reserve margins. The payments would lead to an increase in wholesale prices in the years immediately following their introduction. The payments could then lead to lower wholesale prices because of the construction of additional generating capacity.

Supply Side Scenario #2:
The Power Authority Commits $5 Billion

Workshop participants will discuss the role of the California Consumer Power and Conservation Financing Authority, hereafter referred to as the "Power Authority." The Power Authority was created by SB 6 which calls on the new agency to "achieve an adequate energy reserve capacity in California within five years." SB 6 passed the Senate and was approved by the Governor in May of 2001. The Power Authority was in operation by August of 2001.

Background

The Power Authority has broad powers under SB 6. It could contract with private companies to build plants or it could seize plants owned and operated by private generators. The agency is funded with up to $5 billion in revenue bounds that would be sold to investors and repaid by electricity ratepayers (Contra Costa Times, August 13, 2001). But how should it exercise these broad powers?

S. David Freeman, Chairman of the Board of the Power Authority, is said to favor an activist and enduring role (Contra Costa Times). According to an editorial in the Sacramento Bee, Freeman says his goal is

simply to build enough power plants to provide a 15 percent reserve – anti-blackout insurance, he calls it – that would break the cycle of boom and bust and serve as the backup generator for the state.

Some argue that the Power Authority should act as a "builder of last resort." According to the Contra Costa Times, this is the view of Governor Davis who is said to have "faith in the private sector following through with building a record level of generation." Others say that the Power Authority should follow the example of the New York Power Authority, whose purpose is to "correct an imbalance between supply and demand and provide low-cost power" (Dow Jones Newswires, August 15, 2001).

The creation of the Power Authority has changed the basic structure of the California markets. We now have a hybrid market with a mix of public and private investment. The argument for the public investment can be made in simple, clear terms. According to Freeman (Sacramento Bee):

Somebody has to build power plants. If they don't, we will.

If the Power Authority succeeds in building reserves to 15%, however, some question whether that success will discourage private investment. According to Severin Borenstein, director of the Energy Institute at the University of California at Berkeley,

Once you start an organization like this, it can be very hard to stop it.
It's a real concern that it could turn into the central power player in the California power market.

The Power Authority website describes its challenging mission -- to "ensure sufficient surplus of electricity so that Californians **never again** face electricity shortages." The website argues that the agency can act when necessary to "supplement" private enterprise efforts; it confirms the 15% goal for reserve margins; and it calls for a diverse power portfolio that includes renewable resources. The website reports Letters of Intent for 1,219 MW of wind capacity and 1,846 MW of peaking capacity. It also reports 5,968 MW of peaking proposals that are slated for further evaluation.

The Western Market Model may be used to examine the role of the Power Authority by constructing a scenario that attempts to follow through on the goals of building reserve margins and diversity. The Power Authority's investments will be scheduled during the time period when private investors are reluctant to invest. In this sense, the Power Authority could be viewed as a "builder of last resort" or as a "supplement" to the investments by private developers.

Simulated Impacts

Figure 15 shows Power Authority projects which would use the $5 billion in funding. The scenario envisions 2,000 MW of wind construction starts over the next two years. The model assumes that wind machines cost 1,000 $/kw and require 24 months for construction. Once operational, they are bid into the market as "must run units" and operate with a 33% capacity factor. The wind portion of the Power Authority portfolio would consume $2 billion in capital commitments.

The other $3 billion is committed to 8,000 MW of peakers. The peakers are assumed to be single-cycle, gas burning units which cost 360 $/kw to construct (CEC 2001, p B-9). Their construction interval is quite short, so the model makes them available for operation at the end of the quarter when their construction is started. Their heat rate is 9,300 BTU/kwh. With natural gas priced at 4 $/mmBTU in California, their fuel cost would be around 37 $/mwh. Their variable O&M is 4 $/mwh, so they are bid into the market at around 41 $/mwh.

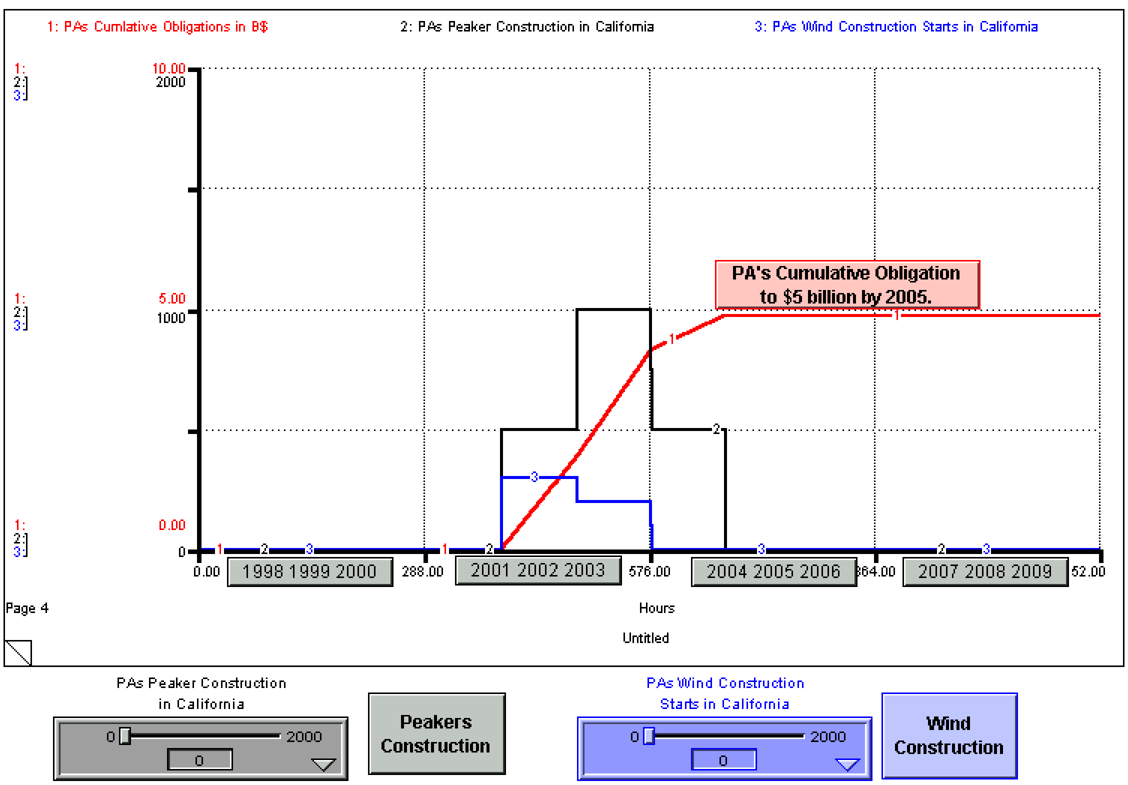


Figure 15. The Power Authority Commits $5 Billion for Peakers and Wind.

Figure 16 shows the reserve margins in the new scenario. These results should be compared with Figure 9 to learn the impact of the Power Authority commitments. This comparison reveals that the Power Authority's actions would allow the minimum reserve margins to reach the 15% target level by the 2003, one year sooner than this goal is achieved in the business as usual scenario. Figure 16 shows that reserve margins would remain above the 15% target during 2004, 2005 and 2006. By 2007, however, the reserve margin falls below 15% and remains below the target for the remainder of the simulation. The Power Authority planners would probably foresee the decline in reserve margins by 2004, and there would be pressure to exceed the $5 billion limit imposed in this scenario.

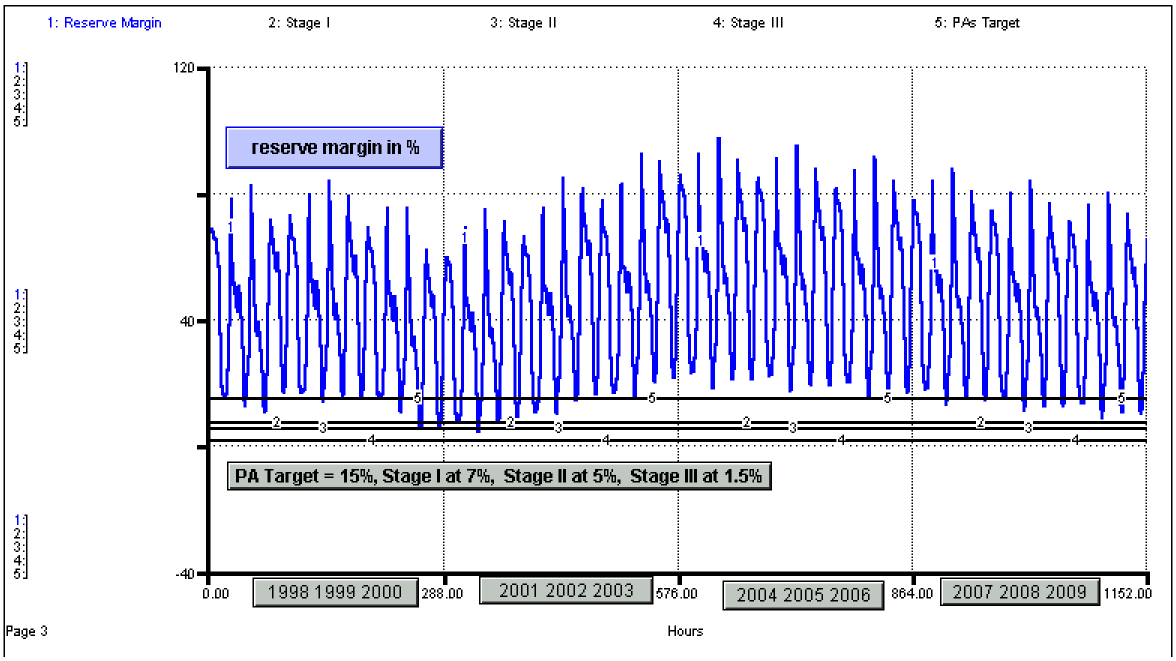


Figure 16. Reserves in the Power Authority Scenario.

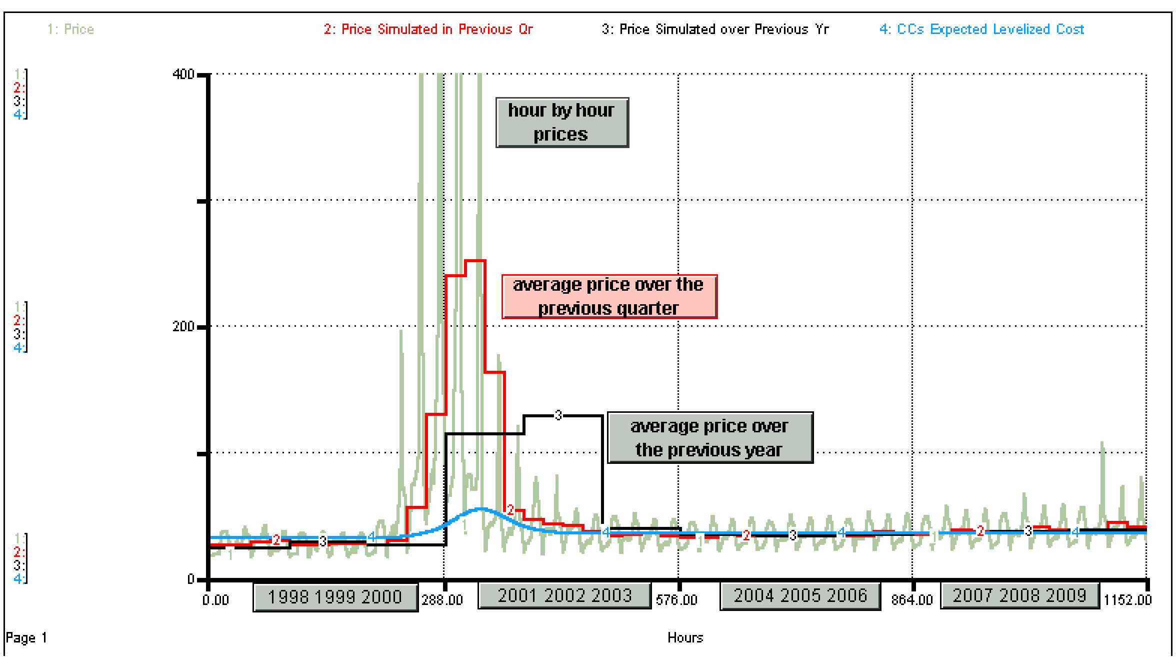


Figure 17. Market Prices in the Power Authority Scenario.

Figure 17 shows the wholesale price implications of the Power Authority Scenario. Comparing these prices with the prices shown in Figure 8 indicates that the Power Authority projects would allow the west to avoid the price spikes that appear in 2006 and 2007 in a business as usual scenario. But there is cause for concern when price spikes reappear in the final year of the simulation. These spikes are another sign that generating capacity is not growing to keep pace with the growth in electricity demand.

Figure 18 shows why generating capacity is not growing as needed in the Power Authority scenario. Much of the new simulation is identical to the previous scenario – private developers are engaged in a major building boom that causes installed CC capacity to grow to around 27,000 MW by the start of the 2004. As before, there is a lull in construction in the years 2004 and 2005. These are the years when private investors are not inclined to build, so the Power Authority could make its commitments and argue that it is acting as the "builder of last resort" or that it is acting to supplement, not to substitute for private investment. The key differences show up in the years 2006 and 2007. During this period, private developers would be hesitant to start new construction. The success of the Power Authority in the previous years has shifted the supply-demand balance to such an extent that private investors do not foresee profitable conditions during the years 2006 and 2007. Rather than start construction, they wait for conditions to improve. As demand grows, conditions improve and a building boom begins around the year 2008. Based on previous simulations, we know that the western system will be vulnerable in the years just prior to the completion of this second building boom.

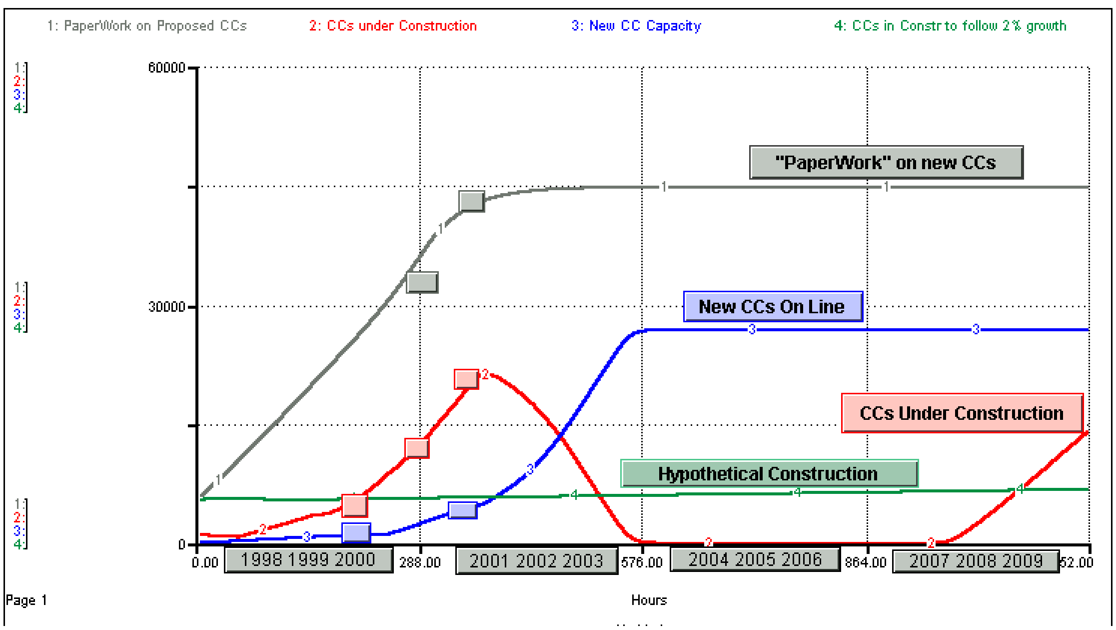


Figure 18. Construction of CCs by private developers
with the Power Authority Committing $5 Billion for Peakers and Wind.

Commentary

This scenario suggests that the Power Authority could make the investments needed to ensure reliable reserve margins and to eliminate a reappearance of price spikes in the western system. By investing in renewable technologies (such as wind) and by timing the investments to occur when private developers are reluctant to build, the Power Authority might appear to be a supplement, not a substitute for private investment. But when the western market is simulated over a longer time interval, it becomes clear that the Power Authority commitments will eventually lead to a reduction in private sector investment. If the Power Authority is to deliver on its mission of ensuring reliable reserve margins, it should be prepared for a large and permanent commitment.

Demand Side Scenario #1: Unfreeze Retail Rates

Background

Retail rates have been frozen under California legislative mandates (AB 1890, AB 265), so electric consumers in California have been somewhat insulated from the dramatic changes in wholesale prices. Several groups have called on California to unfreeze retail rates (CBO 2001, CATO 2001, CERA 2001). They argue that higher rates would help the distribution companies recover the large expenditures made in the wholesale markets on behalf of their retail customers. They also argue that the higher retail rates give the customers a needed incentive to reduce load. Reduced loads could then lead to an improvement in the supply-demand balance and allow the wholesale markets to avoid sky rocketing rates.

Retail rate have not been frozen by legislative rules in the remaining areas of the WSCC. The remaining areas have experienced some retail rate response to the dramatic changes in the wholesale markets. The magnitude and timing of the response has been shaped by a variety of institutions such as the public utility commissions and the federal power marketing agencies. In some instances, retail rates have increased sharply, and there have been significant load reductions in response. In other instances, the threat of large rate increases has led to a negotiated closure of significant industrial loads.

The Western Market Model does not attempt to simulate the wide variety of institutional arrangements for setting retail rates in the west. But it does include a hypothetical distribution company which purchases electricity in the spot market on behalf of its customers. The DisCo is a large company, serving 10% of the retail demand in the WSCC. Its total retail rate was frozen at 87 mills/kwh in all of the previous simulations. This includes a 27mills/kwh charge for generation and a 60 mills/kwh charge for a wide variety of expenses such as distribution, billing, competitive transition charges, and public purposes. The 60mills/kwh charge is held constant and is considered sufficient to cover the DisCo's expenses in all areas except generation.

The generation charge is fixed at 27 mills/kwh, approximating the sales-weighted annual average wholesale price in 1998. The 27 mills/kwh is sufficient to cover the DisCo's spot market purchases in 1998 and 1999. But when the first price spikes appear in the spring of 2000, the DisCo must spend far more in the wholesale market than it recovers from its retail customers. The uncollected expenditures accumulate in a balancing account. The balance in the account climbs rapidly during the year, reaching the staggering value of $15 billion by the end of 2000. In the scenario with unfrozen rates, the generation charge will vary with changes in the wholesale prices. To keep the rate-making rules simple, the model sets the generation charge in the current quarter at the sales-weighted quarterly average price observed in the previous quarter.

Simulated Impacts

Figure 19 shows the total retail rate in the new scenario with the vertical axis scaled from 0 to 250 mills/kwh. The total rate remains close to 87 mills/kwh in 1998 and 1999. There are small variations during the first two years, as the retail rates respond to the changes in the wholesale markets. The large variations appear in the summer and fall of 2000 and in the winter of 2001. The retail rate peaks at 250 mills/kwh, nearly three times higher than the rate changed a year before. (A three-fold increase is similar to the increase temporarily imposed on retail customers in San Diego in the summer of 2000.)

Unfreezing the retail rates provides a tremendous benefit to the cash flow of the DisCo, but the lag in quarterly accounting means that the DisCo will still accumulate significant balances in the balancing account. By the end of 2000, for example, the balance of uncollected expenditures will reach $6 billion. During the course of 2001, the ratemaking rule allows the DisCo to "over-collect" on generation expenditures, and the balance of uncollected expenditures is reduced over time. These results are not shown here (since there is general agreement that the DisCos will benefit from unfrozen rates).

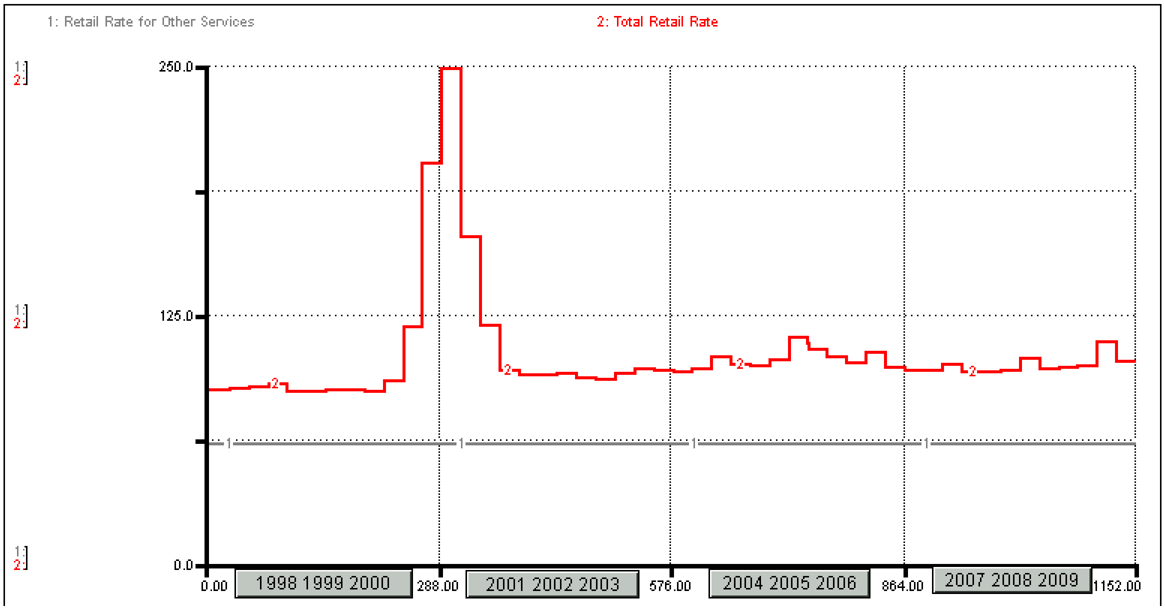


Figure 19. The Retail Rate in Demand Side Scenario #1: Unfreeze Retail Rates.

With the large rate increases shown in Figure 19, consumers would no longer be insulated from the prices in the wholesale market. But how would they respond to this massive rate hike?

The price responsiveness of electricity demand has been studied extensively, especially during the 1970s and 1980s. However, we should remember that very few consumers were exposed to three-fold increases in retail electric rates during the previous decades. Also, it's important to remember that the 1980s and early 1990s was a period of major investments in energy efficiency, so the demand responsiveness may be less today than it was a decade ago.

More recent reports discuss the magnitude of the consumer response to higher rates, but their discussions do not necessarily lead us to an estimate of the price elasticity of demand. A CERA (2001) special report on the California crisis, for example, mentions responses with a price elasticity ranging from as low as 0.03 to as high as 0.60. The CATO Institute, one of the groups advocating higher retail rates, believes that a lower bound on the price elasticity is 0.023, based on the study of San Diego consumers by Bushnell (2001).

To illustrate the impact of unfrozen rates, I have selected an elasticity of 0.2. This is ten times higher than the "extreme lower bound" mentioned by the CATO Institute, but it is at the low end of the range of studies from previous decades. The demand response is implemented as a first order, exponential lag. The length of the lag time corresponds to around two-thirds of the response. The default value is two years based on a review of econometric studies from a previous decade.

Figure 20 shows the simulated consumer response in the form of "Demand Multipliers" on a scale from 0 to 2. The multipliers begin the simulation at 1.0. This means that the demand for electricity is the same as the demand in the previous simulations. The blue curve is the multiplier if there were no lag in the consumer response to higher rates. It drops each quarter, exactly in sync with the increases in retail rates. By the time rates have peaked at 250 mills/kwh, this multiplier has declined to 0.65. If there were no lags in the consumers' response to higher rates, the demands in the WSCC would be 35% below the demands in previous simulations.

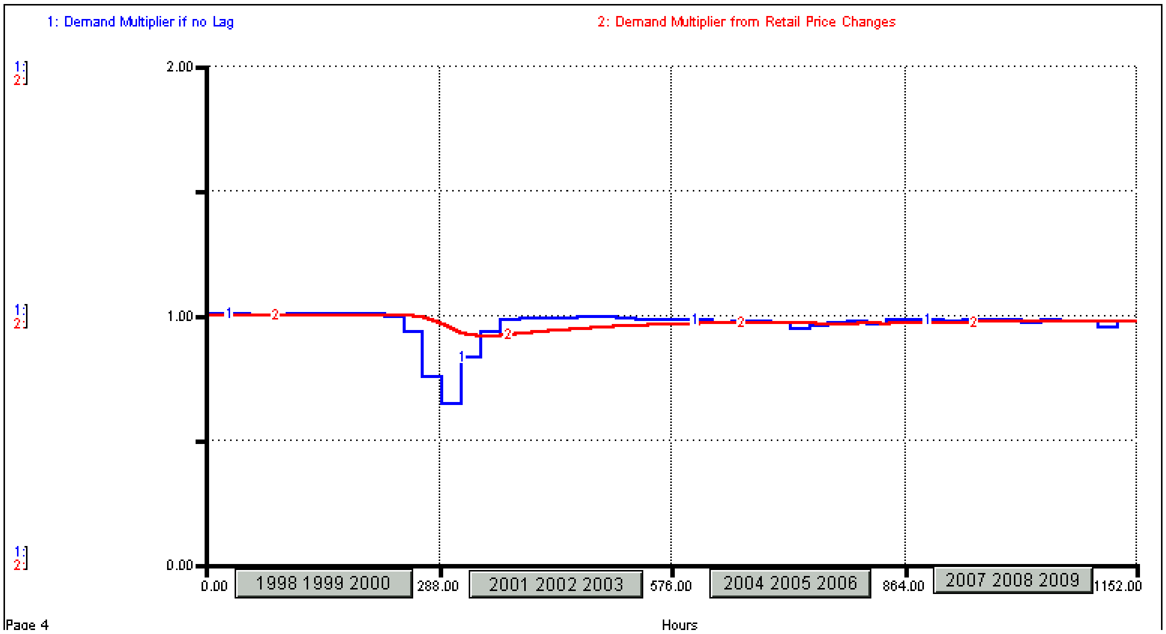


Figure 20. The Customers Respond to Higher Retail Rates.

The red curve shows the actual demand multiplier used to adjust the demands in this simulation. It declines over time, reaching 0.97 by the end of 2000. This 3% response may seem small, but it could prove extremely important to the behavior of the wholesale markets. A 3% reduction across the WSCC lowers the demand by over 4,000 MW. This is an important response, one that is comparable to responses mentioned by those who have argued for retail rate hikes. (The CERA study of the California Crisis states that higher rates would have been extremely helpful if they lowered California demand by around 1,000 to 2,000 MW.)

The red curve in Figure 20 shows that the consumers' response continues well beyond the year 2000. By the end of the year 2001, the multiplier is at 0.92 indicating that consumers have cut their demand by 8% compared to previous simulations. A significant reduction continues in the years 2002 and 2003 as well. (This enduring response makes sense when we think of improved appliances that would be installed in 2000 and 2001 and continue to deliver savings for the remainder of their operating lives.)

Figure 21 shows the demand for electricity in each simulated day, with the vertical axis scaled from 100,000 to 180,000 MW. With this scale, the lower portion of the daily demands is clipped off, so our eye is drawn to the peak demand that appear each quarter. The peak for the year appears in the summer. The summer of 1998 was unusually hot, while the summer of 1999 was unusually cool. This accounts for the noticeable decline in the 1999 peak. The most dramatic decline appears in the year 2001. This is the 8% consumer response to the increase in retail rates. Figure 21 shows that investors are responding to these trends as well.

The model assumes that investors are observing the trends in demand growth over time. Recall from Figure 6 that investors watch the trends in order to prepare a forecast of the peak demand that could appear on the system two years in the future. The demand forecast is combined with their assessment of future supplies to arrive at a price forecast and at a decision on construction starts. When demand begins to fall, the investors will see the decline and adjust their forecasts accordingly. This will lead to a different assessment of the future profitability of a new CC, and the end result could be less construction than in the previous scenarios.

The thin button in Figure 21 is located to help us appreciate the trend in the "Investors' Peak Demand Forecast." The left edge of the button is connected to their forecast prepared in the summer of 2004. The value is around 144,000 MW. This represents their best estimate of the peak demand that might appear in the summer of 2006. The button is two years in width, so its right edge alerts us that 144,000 MW turns out to be reasonably close to the actual demand in the summer of 2006. If you perform similar checks elsewhere in Figure 21, you will see that investors don't always see the future with such accuracy. Their forecast in the summer of 2000, for example, is an overstatement of the peak demand that appears in the summer of 2002. Their forecast in the summer of 2001 errors in the opposite direction. It turns out to be an understatement of the peak demand that appears in the summer of 2003. These errors are entirely plausible since investors can not see the future with "perfect foresight."

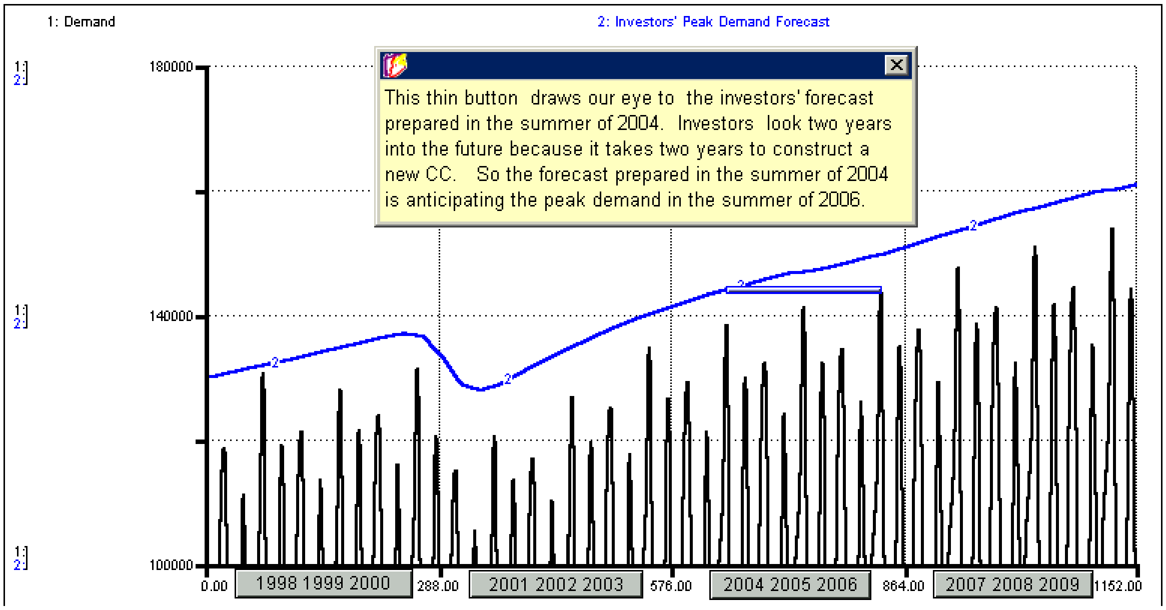


Figure 21. Peak Loads and the Investors' Forecasts of Peak Loads in the Scenario with Unfrozen Rates.

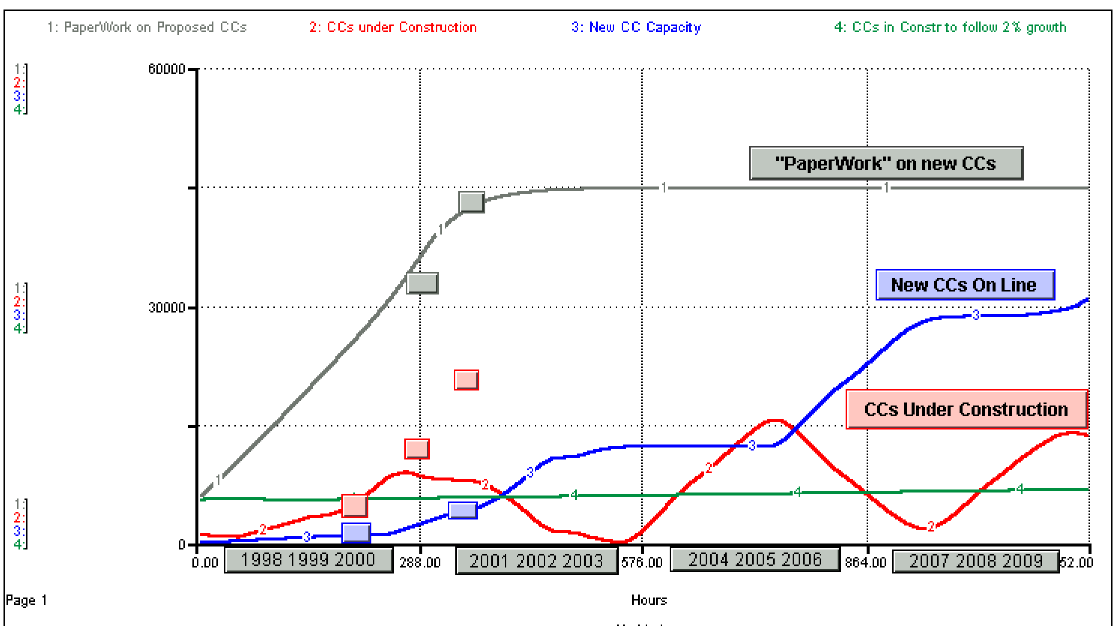


Figure 22. Power Plant Construction in a Scenario with Unfrozen Retail Rates.

Figure 22 shows the power plant construction implications of a scenario with the consumers reacting to higher retail rates. The accumulation of paperwork is the same as in previous simulations, but the actual construction follows a different trajectory. The CCs under construction during 1998, 1999 and early 2000 is similar to previous simulations. Notice, for example, that the red curve comes close to the first of three buttons showing historical benchmarks. But investors foresee a different future near the end of 2000, and the exuberance of the building boom is gone. The total of CCs under construction levels off at around 8,500 MW at the end of 2000. The simulated construction in this scenario is well below the 21,000 MW benchmark for construction midway through 2001. This simulation reveals that major building boom that is underway in the west would not have occurred in a scenario with a significant consumer response to higher retail rates.

Figure 23 shows the wholesale prices in the new scenario. The hourly prices show spikes in the spring, summer and fall of 2000. The summer quarter price turns out to be 123 $/mwh; the fall quarter price is 164 $/mwh. Figure 23 shows a price spike in the winter of 2001, but it is much less severe than we have seen in the previous scenarios. Prices decline during 2001 and remain low during 2002 and 2003.

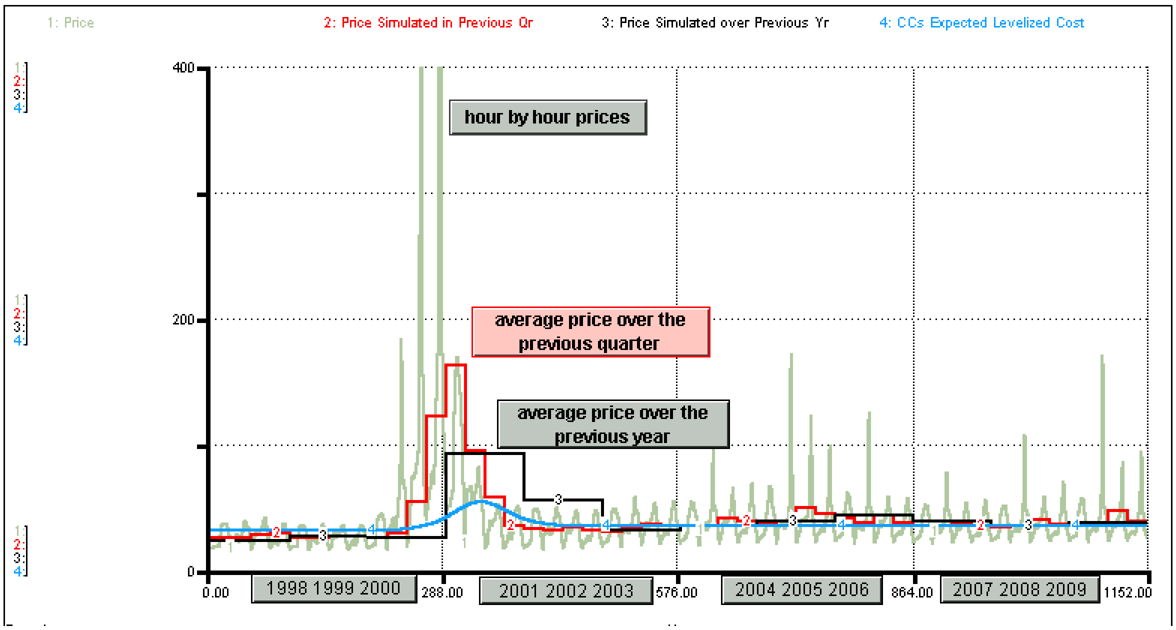


Figure 23. Simulated Wholesale Prices in a Scenario with Unfrozen Retail Rates.

However, price spikes reappear around the year 2004 and 2005 in the new scenario. The spikes arise when generating capacity does not keeping pace with the growth in demand. Figure 22 shows that the investors would be in the midst of a major building boom during 2004 and 2005, but the total installed CC capacity does not increase until around the year 2006. The extra capacity eliminates spikes during the year 2007, but not during the final two years of the simulation.

Figure 24 is a comparative graph showing quarterly average wholesale prices. The business as usual scenario is reported in blue. Wholesale rates exceed 125 $/mwh in the summer of 2000 and are near 250 $/mwh in the fall of 2000 and the winter of 2001. The scenario with unfrozen rates is shown in red. The price comparison shows major benefits in the wholesale markets when consumers respond to the retail rate hikes. Average prices in the fall of 2000 are reduced from 239 to 164 $/mwh, for example. Average prices in the winter of 2001 are reduced from 250 to 95 $/mwh. These are the short-term wholesale rate benefits stressed by those who have argued for higher retail rates.

The second half of Figure 24 reveals the longer-term impact of a scenario with unfrozen rates. Wholesale prices in the new scenario are higher in the years 2004, 2005 and 2006, the years when price spikes appear in Figure 23. For example, the average price for the summer of 2005 is around 50 $/mwh in the new scenario (versus 36 $/mwh in the business as usual scenario). The long-term impact of unfreezing retail rates could be increased prices in the wholesale markets.

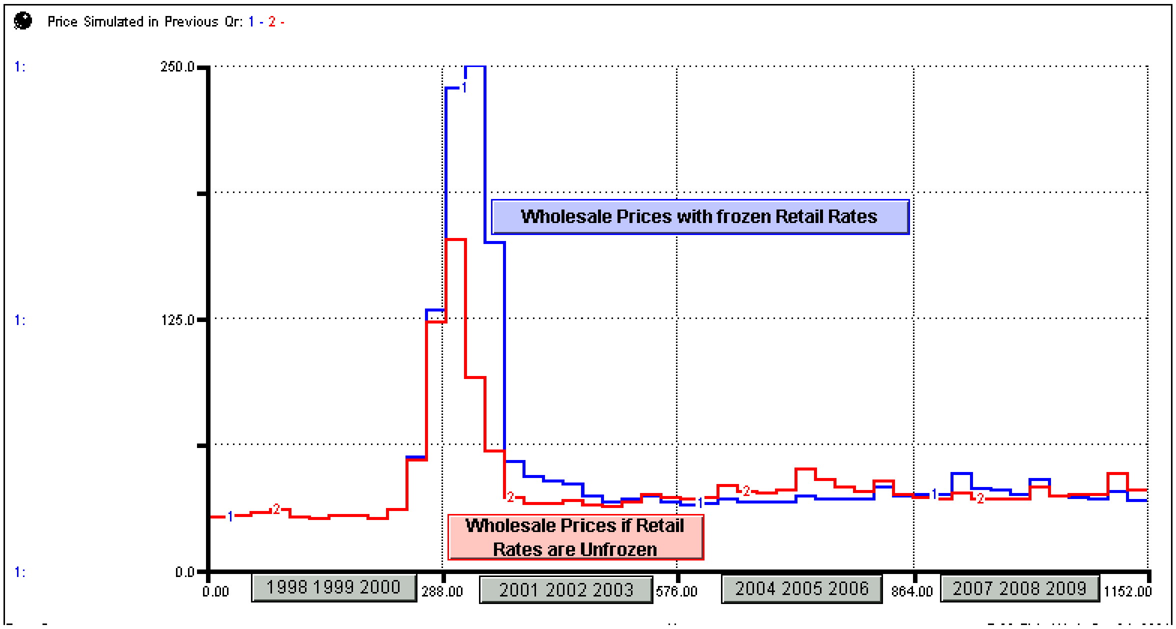


Figure 24. Unfreezing Retail Rates Leads to Lower Wholesale Prices followed by Higher Wholesale Prices

The scenario with unfrozen rates assumes major increase in retail rates, a significant consumer response to the rate hikes, and observant investors who adjust their construction starts as the expected profitability of new CCs changes over time. The resulting scenario shows a new pattern of boom and bust with major changes in wholesale prices during the difficult years of 2000 and 2001. The new scenario then shows rate penalties arising from the new pattern of boom and bust.

These results are somewhat difficult to interpret. Part of the difficulty stems from the fact that a three-fold increase in retail rates did not occur. Interpretation is also made difficult by the extraordinary conditions during 2000 and 2001. This was a period in which small changes in either supply or demand would lead to major changes in the wholesale prices. The scenario with unfrozen rates delivers some important reductions in demand during this crucial time period, and Figure 24 confirms that huge wholesale price benefits that would occur. But the scenario with unfrozen rates also leads to wholesale price penalties later in the simulation. The longer–term penalties are not as dramatic, but we know to expect less dramatic results in the years 2004 and 2005 (because the simulation is not designed for a replay of the extraordinary conditions of 2000 and 2001).

Commentary

I believe unfreezing retail rates will greatly alleviate the distribution company's cash flow and credit problems during the periods of sky rocketing wholesale rates. But unfreezing retail rates does not appear to be effective in breaking the cycle of boom and bust. The scenario with unfrozen rates shows a different pattern of boom and bust, but the fundamental vulnerability would remain. The electric system would be exposed to the possibility of price spikes and outages similar to those shown in previous scenarios.

I concluded earlier in this paper that we could be "one dry year away" from a repeat of the crisis conditions that appeared in the years 2000 and 2001. I draw the same conclusion from the scenario with unfrozen rates. Simply allowing retail rates to respond to a poorly designed wholesale market is not the demand-side answer to the problems of boom and bust.

Demand Side Scenario #2:
Implement Real-Time Pricing

Real-time pricing programs have been proposed to allow customers to participate more directly in wholesale markets. Proponents argue that electricity markets cannot be truly competitive unless both supply and demand participates. Real-time pricing would allow customers to feel the effect of wholesale prices and to reduce load during times of peak prices. These customers would benefit from major reductions in their monthly bills, and the remaining consumers would benefit from the improvement in the supply-demand balance during critical times.

Real-time pricing is also viewed as a method to reduce the exercise of market power. Hirst (2001, p 35) believes that customers who modify their usage in response to wholesale prices provide a "powerful way to discipline the market power that some generators would otherwise have during periods of peak demand." Braithwait (2001, p 52) describes studies of market power in California and in the PJM system and reports that "completely unresponsive demand was one reason cited in both studies for the ability of generators to potentially exercise market power."

Background on Real-Time Pricing

Real-time pricing has been recommended in several reports and papers (Borenstein 2001, Braithwait 2001, CATO 2001, EPRI 2001, Hirst 2001), and the California legislature has allocated $35 million for real-time meters. Programs in California seem to be based largely on a pilot study by Georgia Power Company and some recent experiments in government buildings in California (Levesque 2001). The Georgia Power program is a ten-year pilot study with voluntary participation by 1,600 customers representing around 5,000 MW of peak load. The customers' bills are based on a two-part tariff. They are charged a regulated average rate for a base line load and a real-time rate for additional loads. The program has led to load reductions of around 800 MW during times when wholesale rates are especially high. This amounts to 17% load reductions during peak periods of the most expensive days. The recent measurements of load reductions in California buildings are said to closely match the results from Georgia (Levesque 2001, p. 16).

Braithwait (2001) is impressed by the Georgia results and has argued that real-time prices could be implemented in California, even with average retail rates frozen at legislated levels. He believes the benefits of load reductions during peak hours can be substantial, explaining that industry analysts sometimes estimate a "ten to one" ratio between load reductions and wholesale rate reductions. He uses a wholesale price model (based on cost information from the California ISO) to estimate a 24% reduction in wholesale prices from a 2.5% reduction in load. The "ten to one" result also appears in a discussion of the PJM interconnection by Hirst (2001, p 38) where "a 4 percent drop in demand could have cut the hourly price by almost 50 percent."

Borenstein (2001) describes the potential benefits of two real-time pricing programs that might be implemented in California. One program would be targeted at large customers with 500 kw of load or higher. The metered customers would account for 21% of California peak load or around 8,000 MW. Borenstein gives a conservative estimate of 1,000 MW of load reduction during times of peak prices in the wholesale markets. A more aggressive program would place customers with loads at 200 kw or higher on real-time meters. They would account for around 30% of peak load. Borenstein gives a conservative estimate of 1,500 MW of load reduction.

To illustrate the benefits of real-time pricing, let's consider what might have happened if California had moved to aggressively implement real-time pricing in 1998 and 1999. This scenario will allow us to judge the benefits under the difficult conditions of 2000 and 2001, a time when load reductions would have been extremely valuable.

Simulated Impacts

Imagine what might have happened if we had signed up California's large customers for real-time pricing in 1998 and expanded to the smaller customers in 1999. The MW of RTP load would have grown to 8,000 MW by the start of 1999 and to 12,000 MW by the start of 2000. Furthermore, imagine that this level of price-responsive load is sufficient to discipline the exercise of market power. Figure 25 shows where these assumptions are implemented and viewed in the Western Market Model.

The model assumes that the RTP tariff is designed so that the customers "feel the effect" of wholesale rates in real time, and their response is simulated as the model proceeds through the 24 hours in a typical day. Information on the customers response is quite limited, so it makes sense to proceed with some simple, illustrative assumptions. For example, I assume, that RTP customers will not shed load unless the wholesale price reaches at least 100 \$/mwh. If wholesale prices reach 150 \$/mwh, the load reduction will be similar to the "conservative estimates" by Borenstein in which 12.5% of the participating load will be shed. The load reduction takes place after a one-hour delay for the customers to see the price and for their control equipment to react. If wholesale prices climb even higher, a larger fraction of the participating load will be shed. The maximum response is set at 50%, and this would be triggered by wholesale prices reaching 300 \$/mwh.

This scenario envisions a maximum of 6,000 MW of load reduction if wholesale rates sky rocket like they did in the years 2000 and 2001. Figure 25 shows the simulated load reductions in red. We see the maximum of 6,000 MW appearing in the summer of 2000, followed by somewhat smaller reductions in the fall of 2000 and the winter of 2001. The full 6,000 MW appears again in the spring of 2001. No subsequent reductions appear until the years 2006 and 2007,and these are under 1,500 MW.

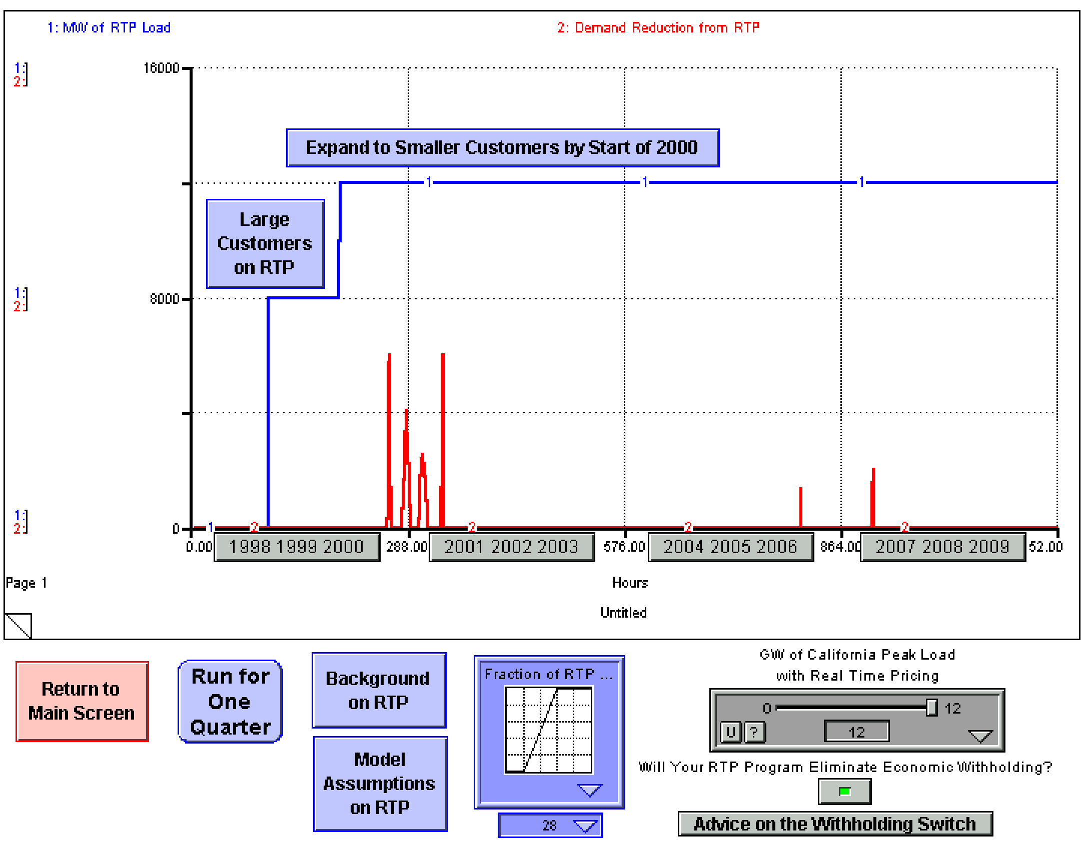


Figure 25. An Aggressive Real-time Pricing Program in California.

As in the previous demand side scenario, it makes sense to assume that investors are watching the implementation of real-time meters and that they are aware of the participating load. But how would investors estimate the fraction of the participating load that might be shed during peak periods in the future? For purposes of illustration, let's assume that they believe around 25% of the participating load would be shed. (This is an arbitrary estimate which is twice as much as Borenstein's "conservative" value of 12.5% but half as much as the maximum of 50%.) With 12,000 MW of RTP load, the 25% assumption translates to a 3,000 MW reduction in the investors' estimates of future peak loads. This reduction is included in their demand assessment (depicted in Figure 6), so we should expect the new simulation to show a different pattern of power plant construction than in the business as usual simulation.

Figure 26 confirms that power plant construction would follow a different trajectory. The CCs under construction remain well below the three buttons which show the construction that has actually taken place during 2000 and 2001. A building boom materializes somewhat later than has actually happened with construction peaking at over 16,000 MW in 2002. This wave of construction would build the installed CC capacity to just over 20,000 MW by the start of 2004. A second wave of construction would appear around 2006-2007, increasing installed CC capacity to around 38,000 MW by the end of the simulation.

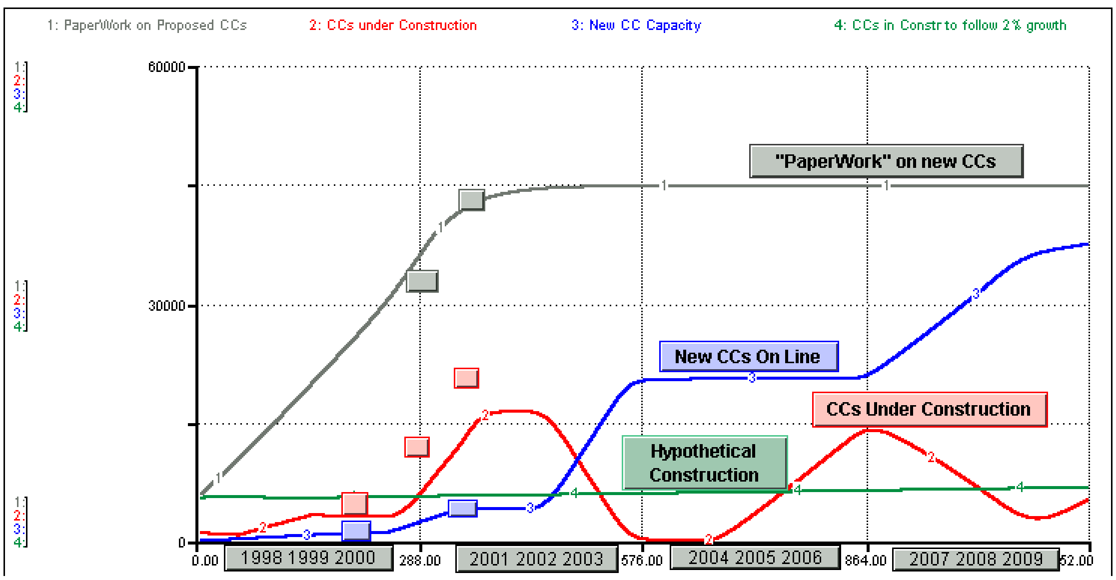


Figure 26. Simulated Construction of New CCs in the Scenario with Real-time Pricing.

Figure 27 shows the wholesale prices in the scenario with 12 GW of RTP load implemented in time for the year 2000. We still see price spikes in the year 2000, but these are to be expected because of the assumption that customers must see the price spikes in "real-time." (Recall that it takes a price spike of at least 100 $/mwh in real-time before RTP customers will shed load.) With this assumption, the RTP program does not eliminate price spikes; it reduces their duration and severity. Figure 27 shows that price spikes reappear in the years 2006 and 2007. These are years when investors are in the midst of the second building boom. The RTP program leads to some load shedding during these years as well, so the duration and severity of the spikes is reduced.

Figure 28 concludes the assessment of the real-time pricing scenario with a comparison chart of wholesale prices in each quarter. The previous simulation shows wholesale prices reaching around 250 $/mwh in the fall of 2000 and the winter of 2001. Real-time pricing shows prices at just under 125 $/mwh. The comparison shows a 50% reduction in wholesale rates from an RTP program with a maximum potential for 6,000 MW of peak load shedding. The 6,000 MW of shedding amounts to around 5% of the peak load in the WSCC. This portion of the comparison confirms the "ten to one" benefits described by Braithwait (2001) and by Hirst (2001).

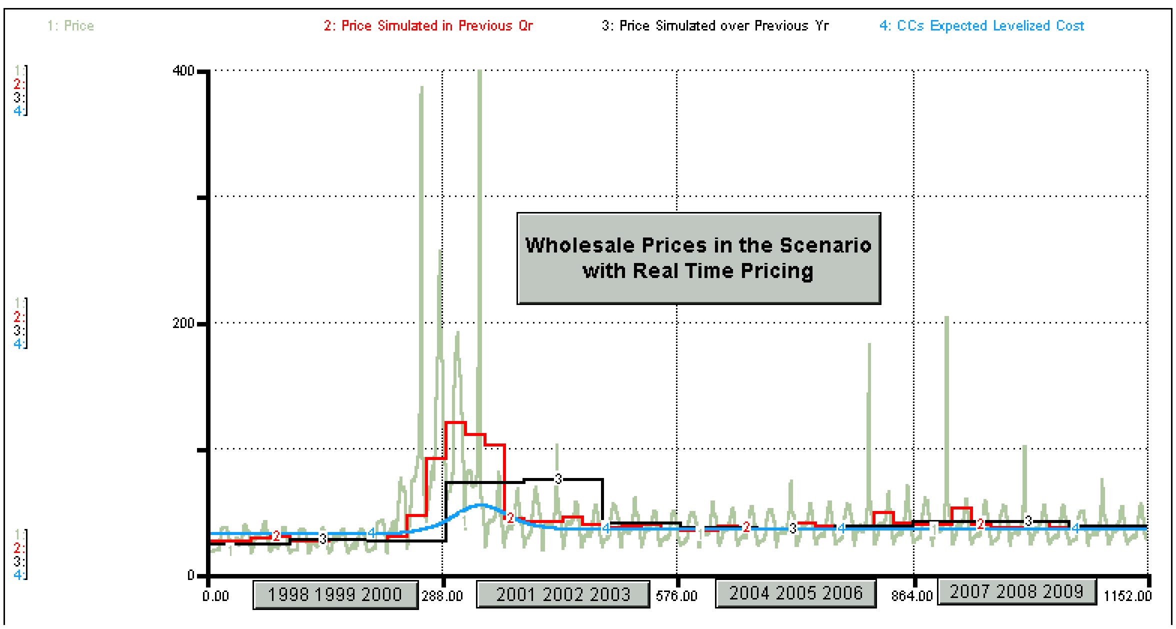


Figure 27. Wholesale Prices in the Scenario with Real-time Pricing.

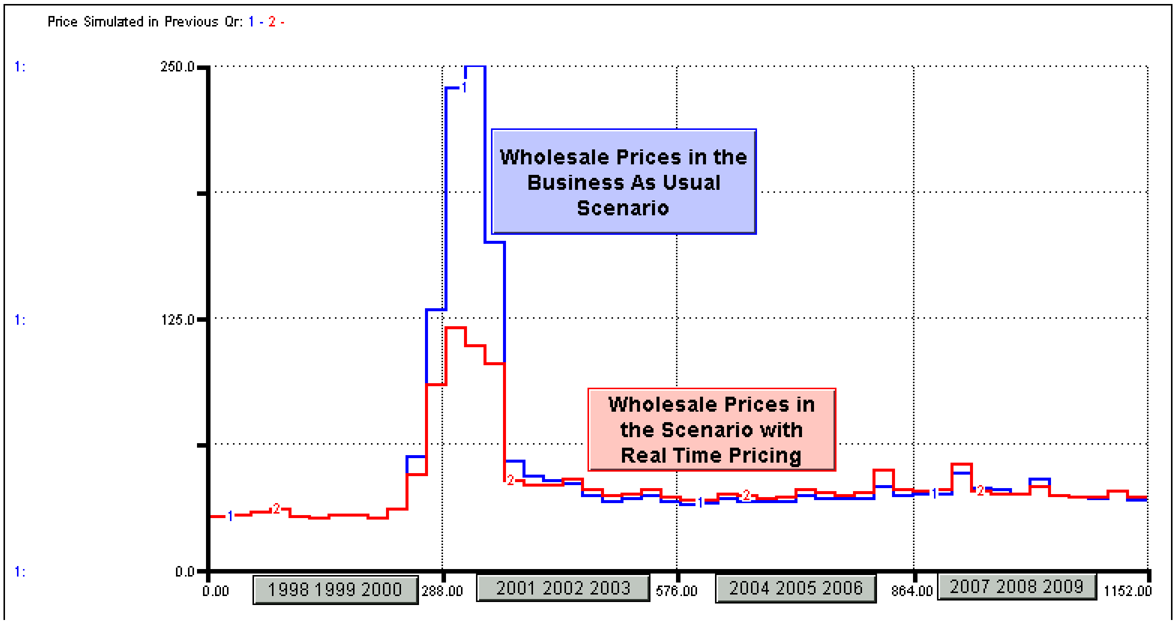


Figure 28. Comparison of Quarterly Average Wholesale Prices.

Figure 28 shows that wholesale prices decline rapidly in both scenarios in the year 2001. The quarterly average prices are rather similar for the remainder of the simulation. Some differences appear in the years 2006 and 2007, a period with price spikes in the real-time pricing scenario. For example, the prices for the summer of 2006 average to 49 $/mwh in the scenario with real-time prices versus 41 $/mwh in the business as usual scenario. These price penalties arise from the reduction in power plant construction in the real-time pricing scenario.

.

Commentary

The real-time pricing scenario is similar in several respects to the scenario with unfrozen retail rates. Both scenarios show important benefits in the wholesale markets during the difficult period of 2000 and 2001. The comparative charts in Figure 24 (unfrozen rates) and Figure 28 (real-time pricing) show comparable benefits from the load reductions that would have appeared had we had more active participation on the demand side. But the demand-side participation is achieved in a much more selective and targeted manner in the real-time pricing scenario.

The two approaches are contrasted below:

- As a method of gaining useful load reductions, unfreezing retail rates is a blunt, unwieldy instrument. In the scenario with unfrozen rates, for example, all of the customers in the WSCC were exposed to a three-fold increase in their retail rates. The price signal was based on the average prices observed over a previous quarter. Many customers are not capable of responding in a significant manner, so they face a three-fold increase in their bills. Others are able to respond, either with behavioral changes on investments in greater efficiency. Some of these reductions will appear in the next quarter, but the majority of the response will be spread further into the future. The end result is load reduction that appear during the bust phase of the boom-bust cycle.

- In contrast, real-time pricing is a precise, targeted instrument for gaining load reductions. It targets the customers with the most capability to respond (through voluntary participation), and it sends signals in real-time. In the simulation scenario shown here, the signal was based on the wholesale price observed in the previous hour. The customers would be equipped with automatic control equipment to allow load shedding within the next hour. The load reductions appear in a timely manner, and they reward the participants with significant reductions in their total bill.

I believe real-time pricing is a more effective and less intrusive means to engage demand-side participation in the wholesale markets. As we implement the meters and tariffs that will allow selected customers to shed loads, we should remember that investors will be aware of the new flexibility on the demand side of the markets. Investors will adjust their assessments of future profitability and postpone investment in new power plants. Based on the simulated scenario shown here, investors would probably continue to build in waves of boom and bust, and the wholesale market would remain exposed to periods of low reserve margins and price spikes.

Appendix A. Closer Look at a Single Day

This appendix shows the "first day" screens of the Western Market Model. These allow for a closer look at the market operations for a typical day in the winter of 1998. Figure A-1 shows the first day results with the hydro generation button in view. It explains that hydro units are dispatched as "must-run" units and that their generation is shaped to provide around a third of the peak demand.

The generations from different technologies in Figure A-1 are shown in a stacked graph, with "other" generation as the second resource in the stack. Other capacity is dispatched as a "must-run" resource. The other capacity is a combination of three WSCC categories (geothermal, other and internal combustion). The WSCC reports over 18,000 MW with over 15,000 MW located in California. Part of the WSCC "other" is wind capacity, but this capacity is removed from the "other category" and dispatched separately (with a 33% capacity factor).

The third resource in the stack is nuclear. The WSCC reports over 9,000 MW of nuclear capacity, almost half in California. The nuclear units are dispatched as "must-run" units with a forced outage rate of 10%. The scheduled outage rate is 20% with the maintenance mainly in the spring.

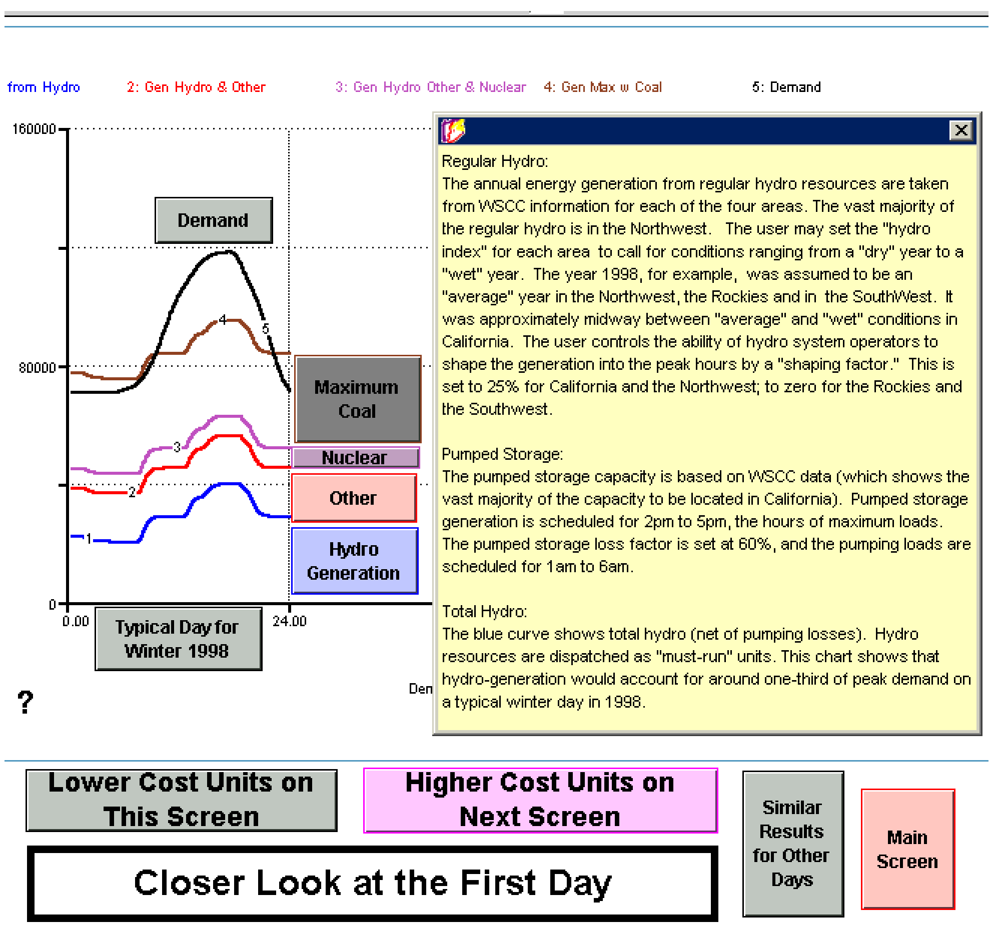


Figure A-1. First Day Results with Hydro Information in View.

Figure A-2 shows the first day results with the coal button in view. It explains that coal-fired units are bid into the market at variable costs and that they would end up serving as the marginal resource during night time hours in a typical day in the winter of 1998.

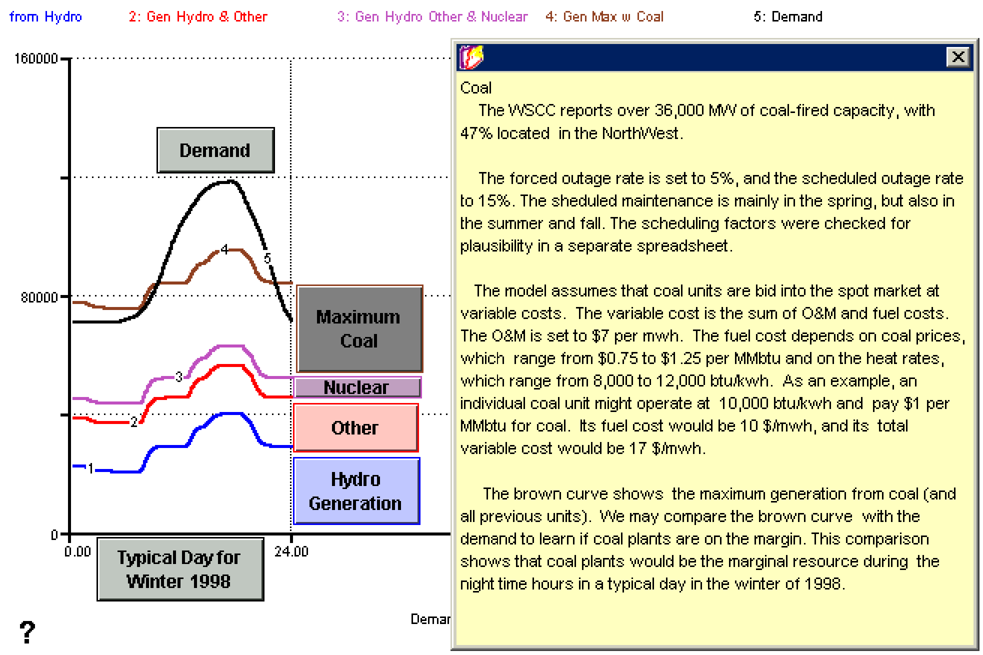


Figure A-2. First Day with Coal Information in View.

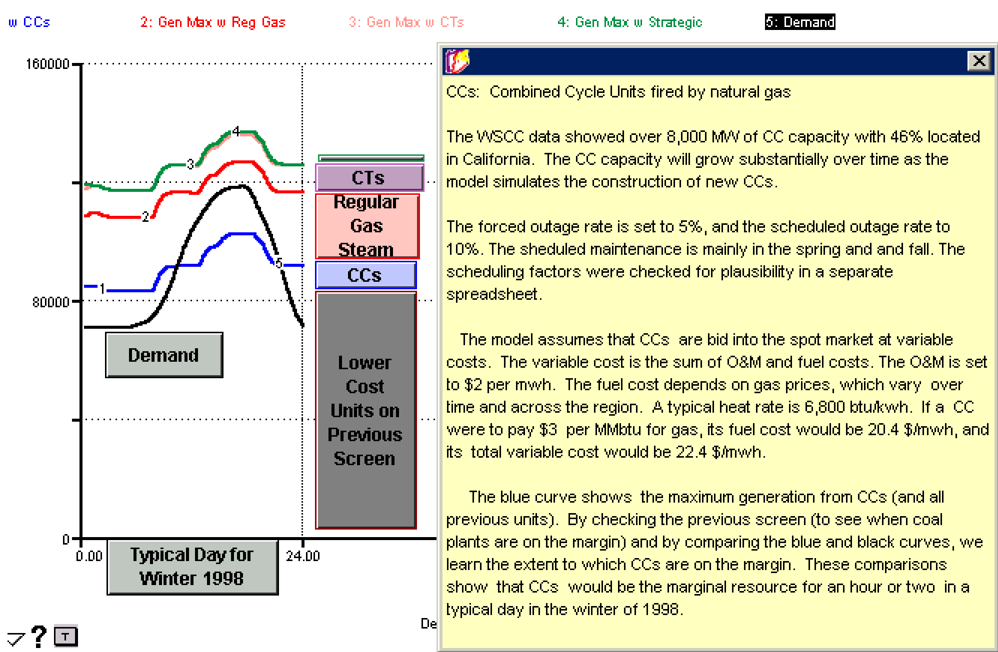


Figure A-3. First Day with CC Information in View.

Figure A-3 shows a different screen with the higher cost resources in view. The information button for the CCs is in view. CCs provide a small portion of the generation in 1998, and they appear on the margin for an hour or two for a typical day in the winter.

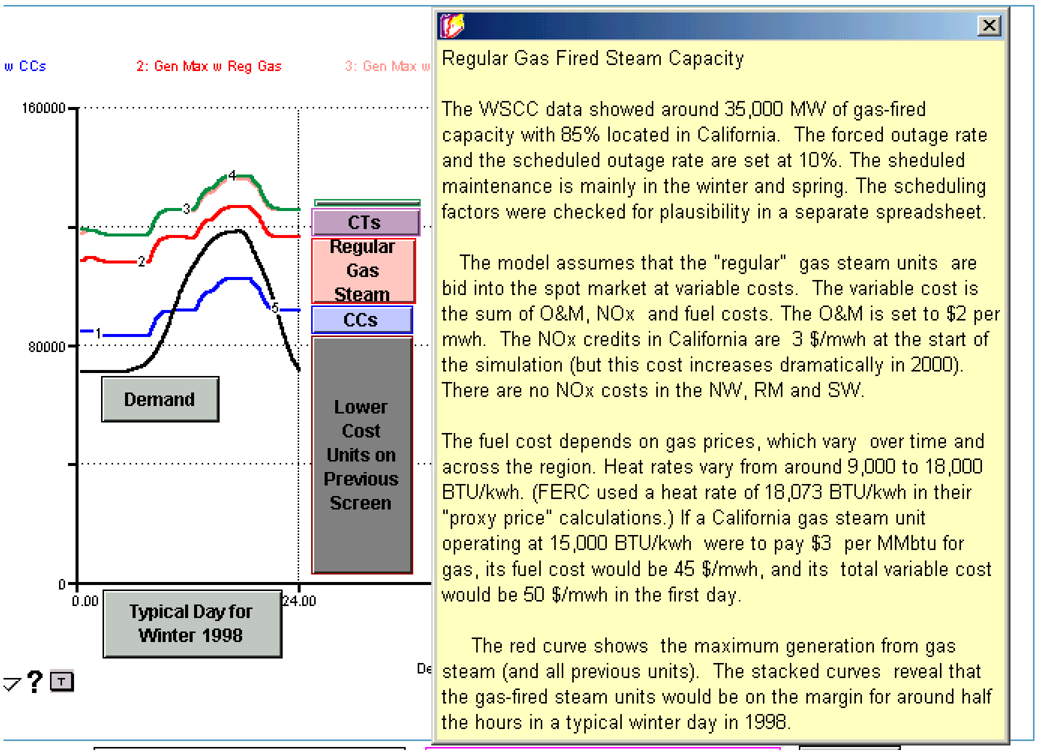


Figure A-4. First Day with Information on the Regular Gas Steam Capacity in View

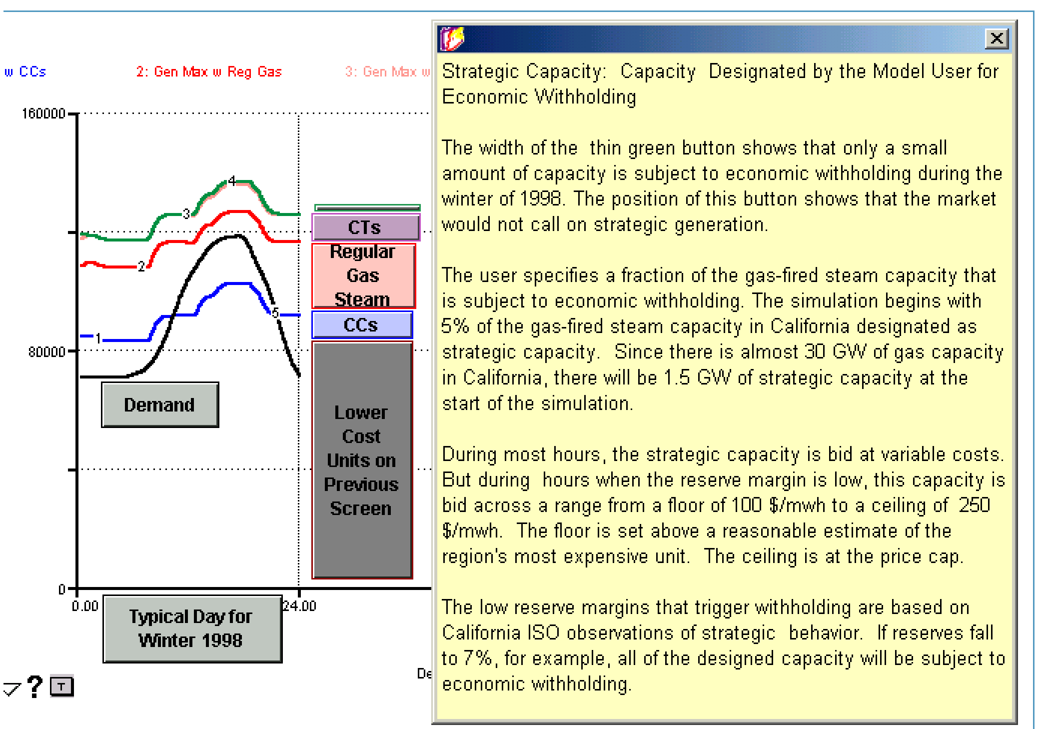


Figure A-5. First Day with Information on the Strategic Capacity in View.

Figure A-4 shows the high cost resources with the gas steam button in view. The "regular" gas steam units are bid into the market at variable costs, and they would appear on the margin for around half the hours in a typical day in the winter of 1998.

CTs are next in the stack. These are single cycle combustion turbines. The WSCC data showed around 9,000 MW of CT capacity, with over 70% in the Northwest and in California. These units are used infrequently, so the model assumes that their forced outage rate and their scheduled outage rate may be set to zero. The model assumes that the CTs are bid into the spot market at variable costs. The variable cost is the sum of O&M and fuel costs. The O&M is set to $3 per mwh. The fuel cost depends on gas prices, which vary over time and across the region. Heat rates vary from around 9,000 to 18,000 BTU/kwh. If a California CT operating at 18,000 BTU/kwh were to pay $3 per MMbtu for gas, its fuel cost would be 54 $/mwh,and its total variable cost would be 57 $/mwh in the first day. Figure A-5 shows that CTs would not appear on the margin in a typical day in winter of 1998.

The final resource in the stack is "strategic capacity." Figure A-5 shows the information on strategic units, gas-fired units which have been designated for economic withholding. Useful information on economic withholding appears in Figure 32 of the California ISO August 10, 2000 Report on California Energy Market: Issues and Performance May-June 2000. The report describes market prices with "competitive outcomes" (usually with total supply in excess of 140% of demand), with "non-competitive" outcomes (with no absolute shortages) and with shortages. The Western Market Model assumes that economic withholding begins when the reserve margin falls below 40% and that a larger and larger fraction of the "strategic capacity" will be withheld as the reserve margin declines toward alert levels. When the reserve margin hits 7%, for example, all of the "strategic capacity" is subject to economic withholding.

Figure A-5 shows no strategic generation for the typical day in the winter of 1998. As demand grows, however, strategic generation will play a larger role. For example, strategic generation is simulated to be especially important in explaining the high prices observed in 2000 and 2001. Further details are given in Appendix B.

Appendix B. Simulated Prices and Actual Prices

This appendix provides a comparison of simulated wholesale prices with actual prices reported by the California ISO. The time axis is expanded to allow a closer look at prices in the first four years. The buttons are located next to the results for each quarter. The ISO information on actual prices is shown in black. The ISO calculates counterfactual prices that would appear if California markets had cleared in a competitive manner. The counterfactual prices are shown in blue. The button for Spring of 2000 is open in Figure B-1. It summarizes the ISO information for April, May and June of 2000. The average price over these months was 78.7 $/mwh. The counterfactual price over these months was 47.3 $/mwh.

The simulated prices from the Western Market Model are shown in red. The average quarterly prices are relatively constant during the first two years and increase dramatically during 2000 and 2001. Figure 3 notes that the first price spike appears in the spring of 2000. This spike contributes to the higher prices for spring of 2000. But Figure B-1 shows that the model falls short of the benchmark. The simulated price is 56 $/mwh, which is around 30% below the actual price reported by the ISO.

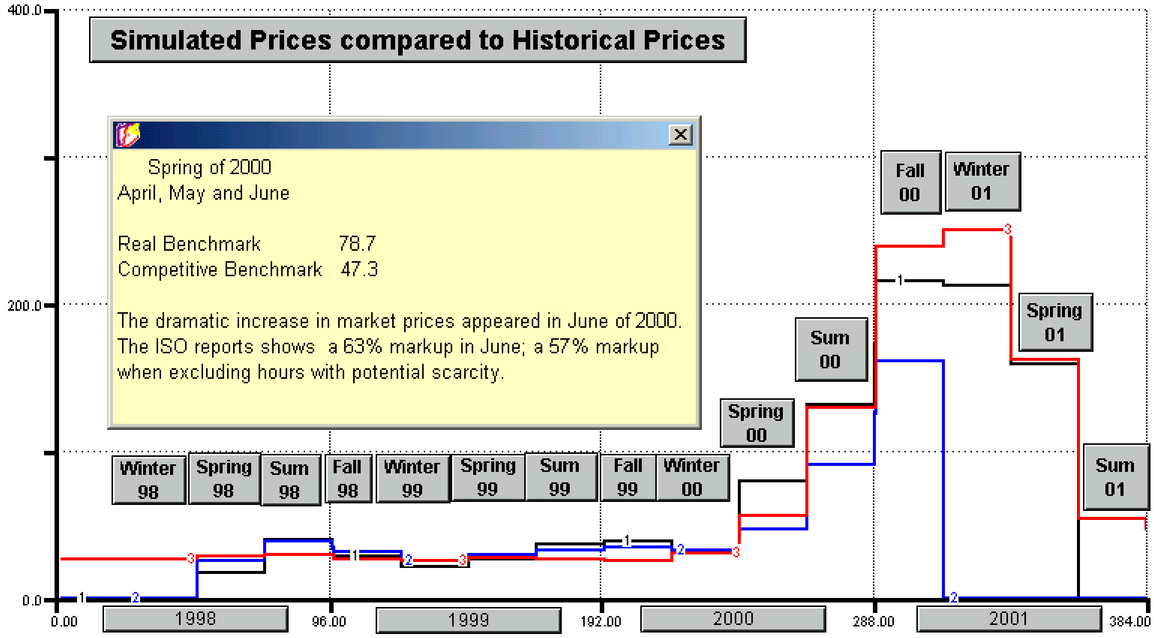


Figure B-1. Comparing Simulated Prices with Actual Prices (Spring 2000 Button Open).

Figure B-2 shows that the model does much better in the summer of 2000. The simulated price (shown in red) is 130 $/mwh which is quite close to the ISO benchmark of 132 $/mwh. The model results are somewhat high in the fall of 2000. The simulation shows the western market clearing at 240 $/mwh, but the ISO reports California markets at 216 $/mwh. The model is somewhat high in the winter of 2001 as well. It shows 251 $/mwh while the ISO reports 213 $/mhw. By spring of 2001, the market is back on target. It gives 163 $/mhw versus 159 $/mwh reported by the ISO. An ISO benchmark for September of 2001 was not available when this display was created. But I did have results for July (63 $/mwh) and August (46 $/mwh). If September prices were similar to August, the average for the summer of 2001 would be 52 $/mwh. The simulation result is 54 $/mwh for the summer of 2001.

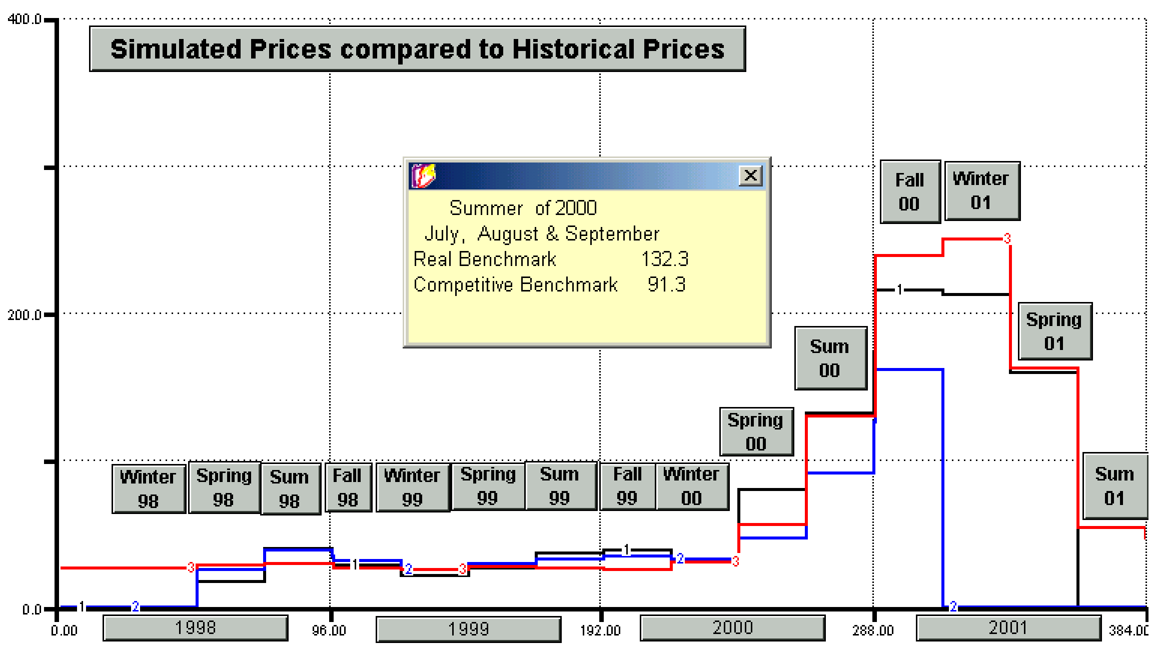


Figure B-2. Comparing Simulated Prices with Actual Prices (Summer 2000 Button Open).

Figure B-3 shows the historical results displayed on the main screen if the user sets the fraction of gas steam capacity for economic withholding to zero. This change allows the model to simulate competitive market prices in the WSCC system. The market prices are displayed in the same format as Figure 3. Recall from Figure 3 that the first price spike appeared in the spring of 2000, peaking at nearly 200 $/mwh. The button in Figure B-3 draws our attention to a similar price spike in the new simulation. We learn that price spikes could still appear with competitive assumptions, but they would not be as severe.

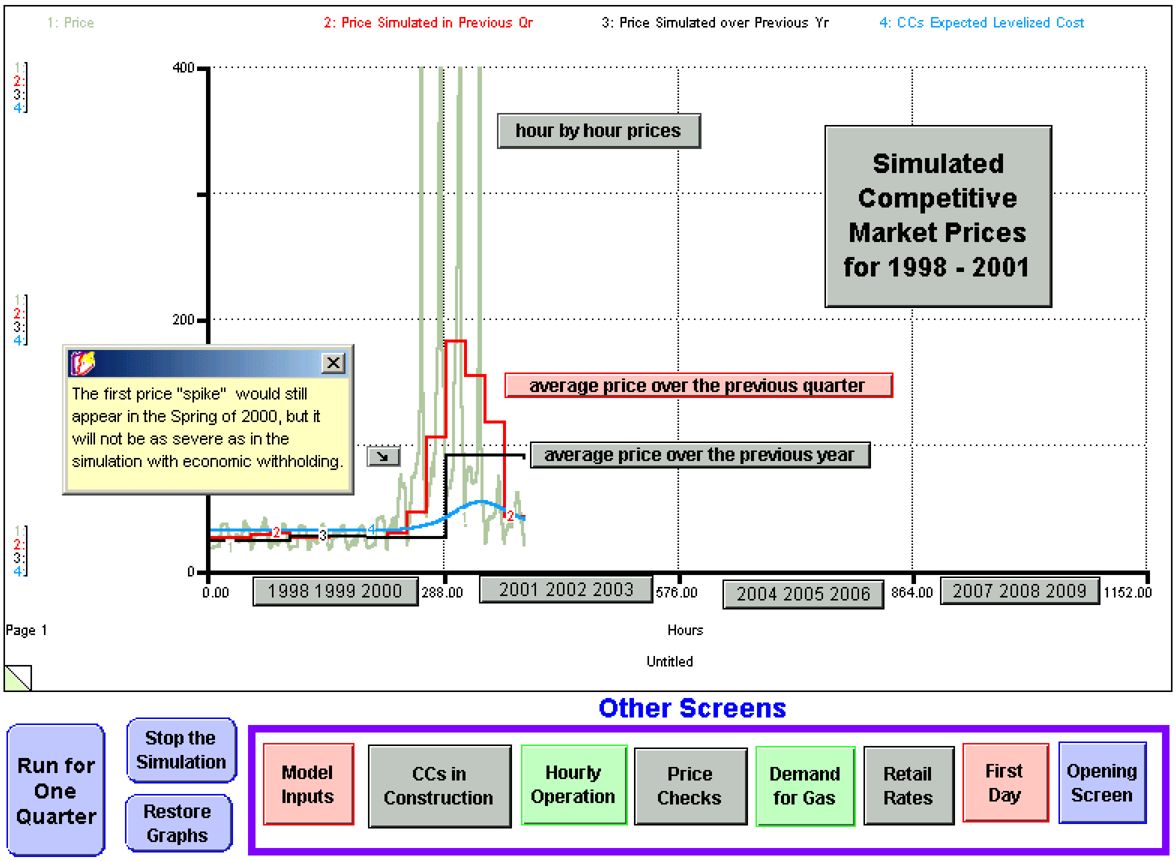


Figure B-3. Main Screen with four Years of Simulated Prices if there is no economic withholding.

Figure B-4 shows the price checks screen with simulated prices arranged for comparison with the ISO counterfactual prices. As before, the model results are shown in red. But in this case, we compare the model results with the ISO counterfactual prices shown in blue. The spring 2000 button is open, so we can read that the ISO benchmark was 47.3 $/mwh. The simulation result comes quite close at 46.5 $/mwh. The simulation result is somewhat higher than the ISO counterfactual result in summer of 2000 (106 versus 91 $/mwh) and in the fall of 2000 (182 versus 161 $/mwh). I did not have ISO counterfactual results in 2001 when this display was created, so the blue curve drops to zero at the end of the year 2000.

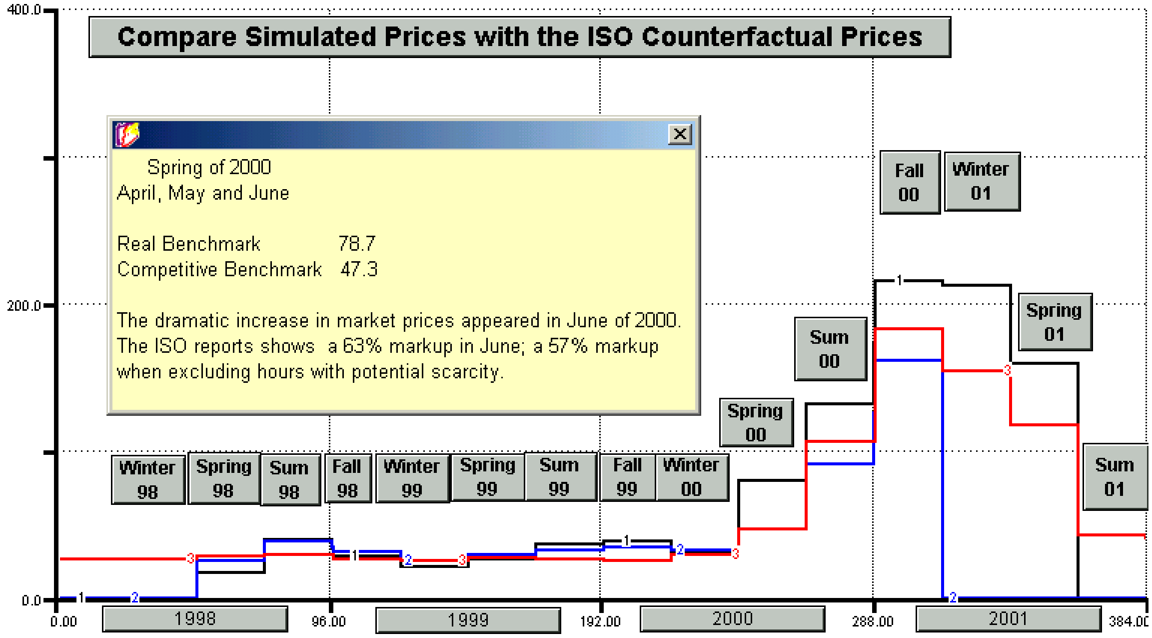


Figure B-4. Comparing Simulated Prices with ISO Counterfactual Prices.

Appendix C. The Size of the Current Building Boom

Figure 7 shows a business as usual scenario in which CCs under construction peaks near the end of 2001. The current building boom leads installed CC capacity to grow to around 27,000 MW by the start of 2004. The scenario shows a two-year lull in construction followed by a second wave of construction. By the end of the simulation, there is 43,000 MW of installed CC capacity in the WSCC.

This pattern of boom and bust construction arises from the theory of investor behavior shown Figure 6 and implemented with the inputs shown in Figure C-1. This screen shows the attributes of new CCs with a total levelized cost around 31 $/mwh near the start of the simulation. The CCs require 12 months for permitting and 24 months for construction. The equations for developers' permits builds the paperwork toward the 45,000 MW target shown in Figure 7. The permits are distributed among the four areas of the WSCC based on the distribution of power plant construction in a recent review. Investors look into the future to anticipate the growth in demand, the amount of generating capacity and the system reserve margin. From this information, they estimate the average market price over the course of the year shortly after a new CC would enter operation. The investors watch current generating capacity. As soon as units are retired, they adjust their forecast downward. (They do not try to guess when future retirements might occur.)

This appendix focuses on the weight given to capacity in the pipeline. The "weight" in Figure C-1 represents how the investors account for the CCs under construction in forming their estimates of future market prices. We set the weight to 1 if we believe that all of the investors count the CCs in their forecasts while they are still under construction. We set the weight to 0 if we think none of the investors count these units in their forecasting. The evidence from industries like commercial real estate suggests that investors are not inclined to count all of the capacity in the construction pipeline in their decision-making. Figure C-1 shows the base case assumption at 0.5, an intermediate value which gives results similar to the dominant pattern published previously (Ford 2001).

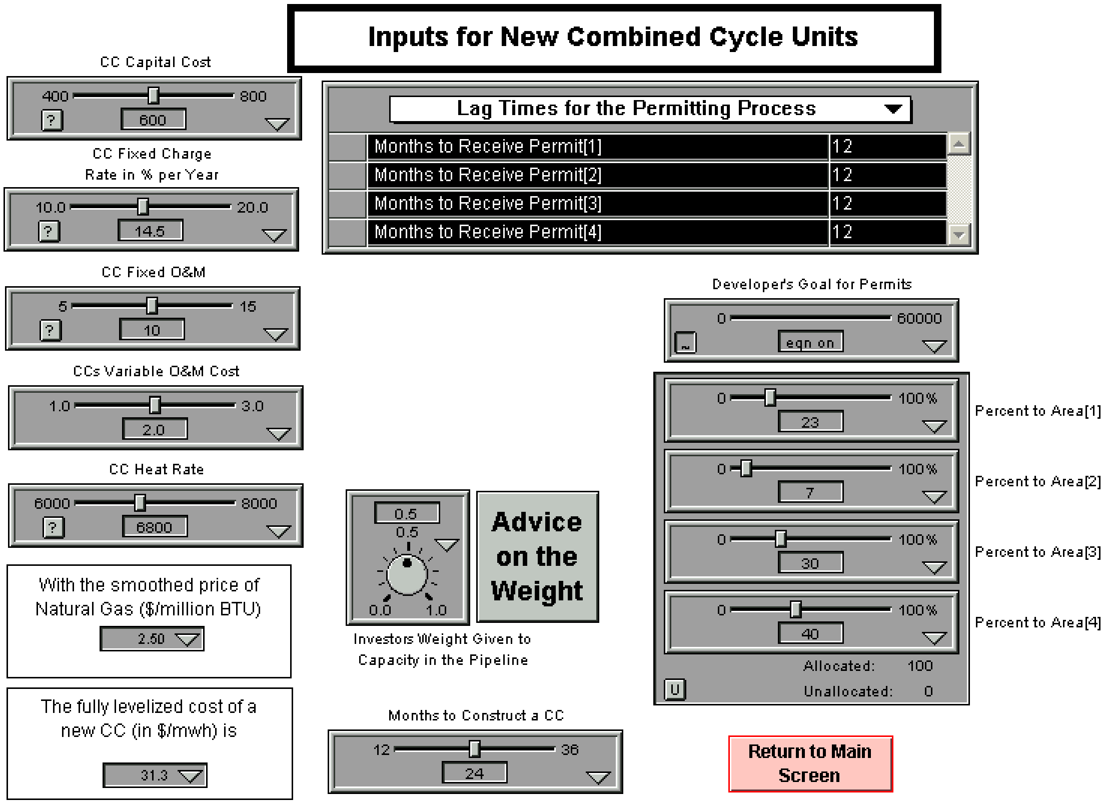


Figure C-1. User Inputs for New Combined Cycle Power Plants.

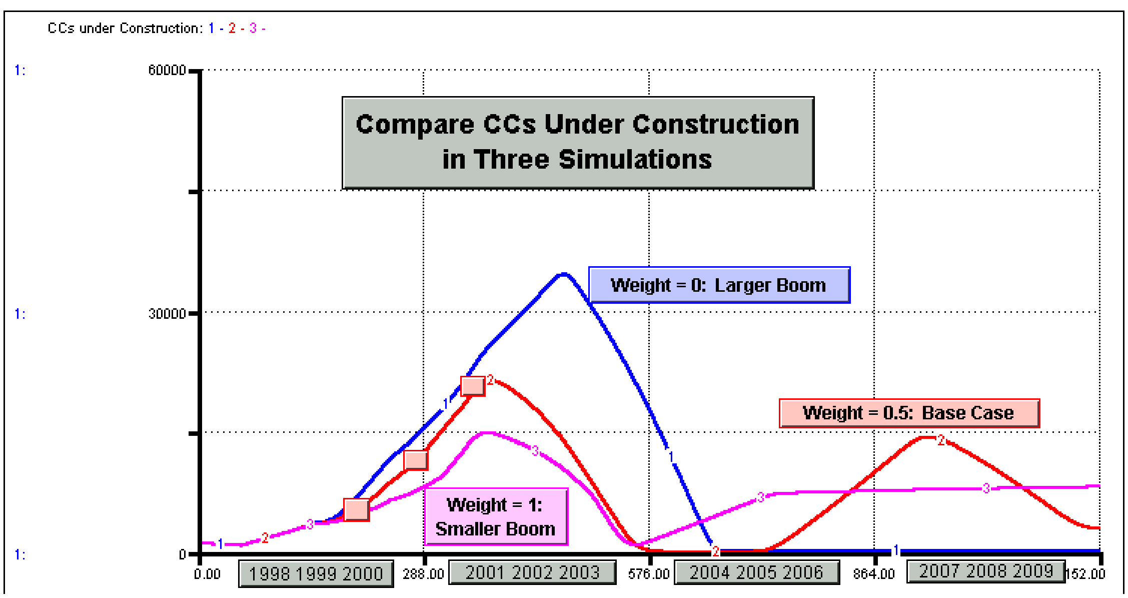


Figure C-2. Comparison of CCs Under Construction with in Three Simulations.

Figure C-2 compares the CCs under construction in three simulations with different weights assigned to the CCs under construction. With the weight set to 0, we see a more prolonged building boom which does not crest until the end of 2002. CC capacity would grow to around 51,000 GW by the end of this building boom. This huge building boom would build reserve margins to well above 15% for the simulation. Figure C-2 shows that this simulation generates construction activity slightly higher than the historical benchmarks. The results with the weight at 0.5 have been shown previously. This is the business as usual scenario reported in Figure 7. Figure C-2 shows that the base case provides a somewhat better fit with the historical benchmarks. CC capacity would grow to around 27,000 MW by the end of the first building boom.

The third simulation shows the pattern of construction if investors counted 100% of the CCs under construction in their decision-making. This assumption leads to a smaller boom, one that falls well short of the historical benchmarks at the end of 2000 and midway in 2001. If we believed a weight of 1.0 were realistic, however, we would expect to see investors adjust their construction shortly after the first building boom and arrive at a dynamic equilibrium with a relatively constant amount of construction activity year after year. This level of construction turns out to be insufficient to deliver a reliable system. Reserve margins (shown on a different screen) would trigger alerts from 2005 until the end of the simulation. Price spikes (shown on another screen) would appear by the year 2005 and remain for the rest of the simulation. These results are similar to patterns published previously (Ford 2001).

Appendix D. The Demand for Gas for Power Generation

The dramatic increases in natural gas prices is one of the factors contributing to the unusually high prices in the western electricity markets in the year 2000. The natural gas system has been examined in recent studies by the State of California (CEC2001) and the State of Washington (WA OTED 2001). These studies remind us of the important linkages between the natural gas supply and the competitive electricity markets.

The Western Market Model accepts user inputs for gas prices in each of the four WSCC areas as user inputs. The simulations shown in this report set the long-term price of gas at 4 $/mmBTU in California. (According to an August 17, 2001 article in the San Jose Mercury News, this value is in the middle of a range of values used by officials at the Pacific Gas and Electric Company.) The price of gas in the rest of the WSCC is set at 3 $/mmBTU. These prices are fixed, regardless of the amount of gas that is consumed in the power sector. This appendix reports the amount of gas that would be consumed in power generation in the WSCC. The gas consumption patterns will provide some context for evaluating the assumption that gas prices will remain constant in the coming years.

Gas consumption is based on gas-fired generation (see Appendix A) and the average heat rates of the units that are in operation. Figure D-1 shows the gas consumption as the model simulates a typical day in each quarter. The hour by hour consumption is shown in gray on a scale from 0 to 8 MMcf per hour. The results for each 24 hour day are accumulated and then reported for the quarter on a scale from 0 to 8 Bdf per day. The most dramatic result in Figure D-1 is the huge increase in gas consumption during the summer of 2000. Many believe that this increase contributed to the tight supply conditions in the year 2000 and is one of the reasons for the sharp increase in gas prices in the year 2000.

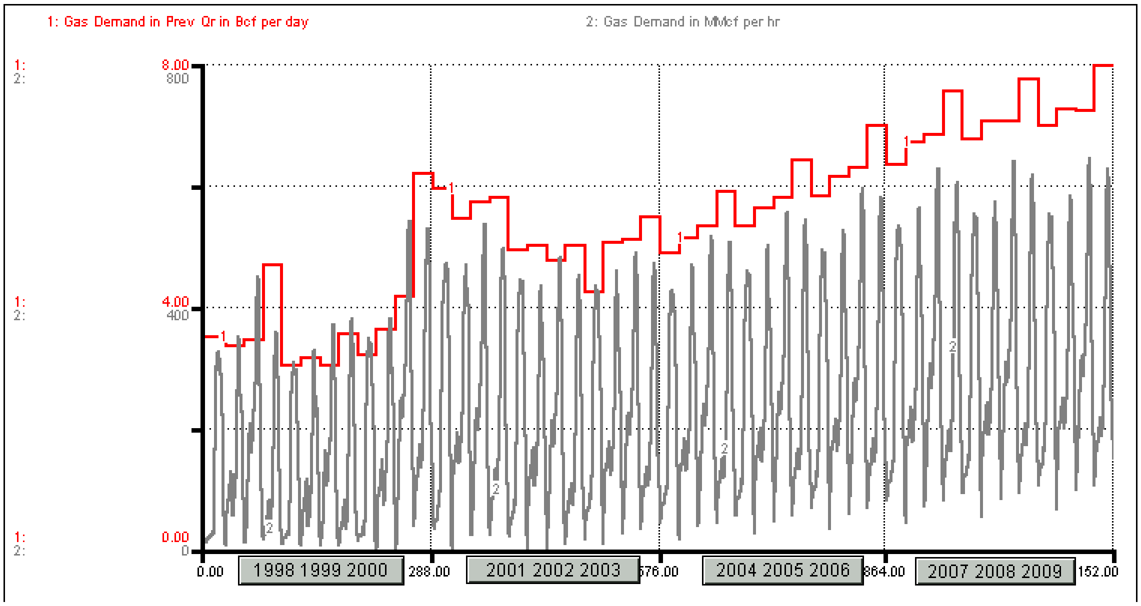


Figure D-1. Demand for Natural Gas in Power Generation in the Business as Usual Scenario.

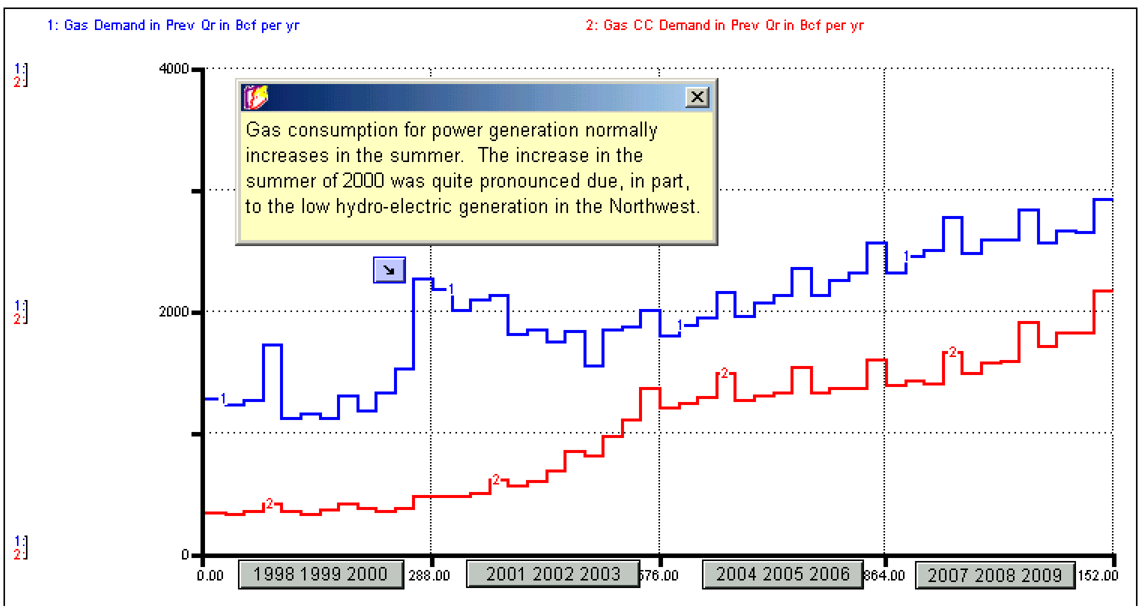


Figure D-2. Total Gas Demand and the Demand from Combined Cycle Units.

The blue curve in Figure D-2 shows the total gas consumption for each quarter reported in Bcf per year. The State of Washington (WA OTED 2001) reports gas consumption in the west at 1,198 BCF for 1999 and at 1,761 BCF for 2000. The model results for 1999 are 2% below the WA OTED estimate; the results for 2000 are 2% high.

The red curve in Figure D-2 shows the gas consumption from the CCs in the same units and on the same scale. The construction boom leads to a major increase in installed CC capacity in the years 2002 and 2003. Figure D-2 shows that the gas consumption from these new units would increase dramatically during the years 2002 and 2003. But total gas consumption is shown to decline during this time period. These are the “bust years” in the boom/bust cycle shown in Figure 8. The system operator is able to meet electricity demands without accepting the bids of the older, less efficient gas-burning units. Older units would not be operated as extensively during these years, and total gas demand remains relatively flat for three or four years. After 2004, the total demand begins to grow again. We know from Figure 7 that installed CC capacity is constant at this time, so the growth in gas demand is coming from greater operation of the older, less efficient gas-burning units.

Figure D-3 compares the quarterly gas demands in two simulations. The base case simulation is shown in blue. The alternative case assumes a much more vigorous building boom (described in the previous appendix). The investors are assumed to discount CCs currently under construction, and the current building boom does not crest until the year 2003. Figure D-3 shows that a more exuberant building boom would actually lead to lower consumption of gas in the western electric system. The reduction in gas consumption arises from the reduction in market clearing prices that appear in the simulation with a larger building boom. Lower market prices lead to lower operation of the region's older, less efficient gas-burning units.

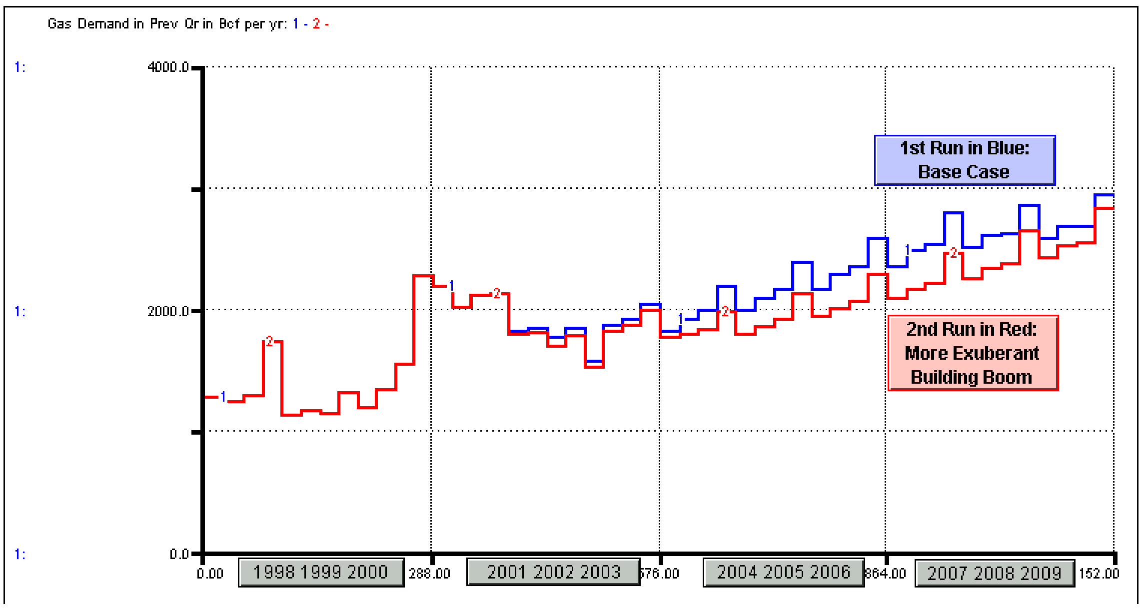


Figure D-3. Comparison of Gas Consumption for Power Generation in Two Simulations.

References

Borenstein 2001: Severin Borenstein, Frequently Asked Questions About Implementing Real-time Pricing in California for Summer 2001, University of California Energy Institute, March 2001.

Braithwait and Faruqui 2001: Steven Braithwait and Ahmad Faruqui, The Choice Not to Buy, Public Utilities Fortnightly, March 15, 2001.

Bunn and Larsen 1992: Derek Bunn and Erik Larsen, Sensitivity of Reserve Margin to Factors Influencing Investment Behaviour in the Electricity Market of England and Wales, Energy Policy.

Bushnell 2001: James Bushnell and Erin Mansur, The Impact of Retail Rate Deregulation on Electricity Consumption in San Diego, working paper PWP-082, University of California Energy Institute.

CATO 2001: Jerry Taylor and Peter VanDoren, California's Electricity Crisis, policy analysis paper no. 406 from the Cato Institute, July 3, 2001.

CBO 2001: Congressional Budget Office, Causes and Lessons of the California Electricity Crisis, Sept 2001.

CEC 2000: California Energy Commission, Market Clearing Prices Under Alternative Resource Scenarios 2000 – 2010, report P300-00-001, Feb 2000.

CEC 2001: California Energy Commission, Natural Gas Infrastructure Issues, Staff Draft Report P200-01-001, May 2001.

CEC 2001: California Energy Commission, California Energy Outlook—Electricity and Natural Gas Trends Report, Staff Draft of Report 200-01-002, Sept 2001.

CERA 2001: Cambridge Energy Research Associates, Special Report—Beyond California's Power Crisis: Impact, Solutions, and Lessons, March 1, 2001.

Dailey 2001: Will Dailey, Supply Side Boom, Public Utilities Fortnightly, Oct 15, 2001.

DOE 2001: Office of Policy, U.S. Department of Energy, The Impact of Wholesale Electricity Price Controls on California Summer Reliability, June 2001.

DiPasquale 1996: Denise DiPasquale and William Wheaton, Urban Economics and Real Estate Markets, Prentice Hall.

Dyner and Larsen 2001: Isaac Dyner and Erik Lasrsen, From Planning to Strategy in the Electricity Industry, Energy Policy, Vol 29, p. 1145-1154.

EPRI 2000: F. Graves, A. Ford and S. Thumb, Prospects for Boom/Bust in the US Electric Power Industry, Technical Report 1000635 of the Electric Power Research Institute, December 2000.

EPRI 2001:The Western States Power Crisis: Imperatives and Opportunities, an EPRI White Paper, June 25, 2001.

EPRI 2001: Ahmad Faruqui, Hung-po Chao, Vic Niemeyer, Jeremy Platt and Karl Stahlkopf, Analysing California's Power Crisis, The Energy Journal, Vol 22, No. 4

FERC 2001: Federal Energy Regulatory Commission, June 19, 2001 Order on rehearing of monitoring and mitigation plan for the California wholesale electric markets.

Ford 1999: Andrew Ford, Cycles in Competitive Electricity Markets, Energy Policy, Vol 27, p 637-658.

Ford 1999: Andrew Ford, Modeling the Environment, Island Press.

Ford 2001: Andrew Ford, Waiting for the Boom: A Simulation Study of Power Plant Construction in California, Energy Policy, Vol 29, p 847-869.

Forrester 1961: Jay W. Forrester, Industrial Dynamics, Pegasus Communications, Waltham, MA.

Hildebrandt 2001: Eric Hildebrandt, Dept. of Market Analysis, California ISO, Further Analyses of the Exercise and Cost Impacts of Market Power in California's Wholesale Energy Market, March 2001.

Hirst 2001: Eric Hirst, Price-Responsive Retail Demand: Key to Competitive Electricity Markets, Public Utilities Fortnightly, March 1, 2001.

Hoyt 1933: Homer Hoyt, One Hundred Years of Land Values in Chicago, University of Chicago Press.

Hunt 1996: S. Hunt and G. Shuttleworth, Competition and Choice in Electricity, Wiley, New York.

ISO DMA 2000: California ISO, Dept. of Market Analysis, Report on California Energy Market Issues and Performance: May-June 2000, August 10, 2000.

ISO MSC 1999: California ISO Market Surveillance Committee, Report on Redesign of California Real-time Energy and Ancillary Services Markets, October 18, 1999.

ISO MSC 1999: California ISO Market Surveillance Committee, Annual Report on Market Issues and Performance, June 1999.

Kahn 2000: Edward Kahn and Paul Joskow, A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Markets During Summer 2000, FERC Testimony, Nov 21, 2000.

McCullough 1998: Robert McCullough, Can Electricity Markets Work Without Capacity Prices? Public Utilities Fortnightly, March 15, 1998.

Michaels 1997: R. Michaels, MW Gamble: The Missing Market for Capacity, The Electricity Journal.

Levesque 2001: Carl Levesque, Electricity Evolution in the UK, Public Utilities Fortnightly, July 15, 2001.

Levesque 2001: Carl Levesque, Real-time Metering: Still as Sweet with Prices Controlled, Public Utilities Fortnightly, Sept 1, 2001.

Newbery 1995: D. Newbery, Power Markets and Market Power, The Energy Journal, 16 (3).

NPPC 2000: Northwest Power Planning Council, Study of Western Power Market Prices, Final Report, October 11, 2000

Oliver 2000: Testimony of Stephen Oliver, VP Power Business Line of the Bonneville Power Administration before the Senate Energy and Natural Resources Subcommittee Hearings on Northwest Electricity Prices, Oct 5, 2000.

PJM MMU 2000: PJM Market Monitoring Unit, Monitoring the PJM Energy Market: Summer 1999, Proceedings of the Fifth Annual Research Conference on Electricity Industry Restructuring, University of California Energy Institute

Power Authority 2001: California Power and Conservation Financing Authority website, http://www.capowerauthority.ca.gov/background/main.asp viewed on October 8, 2001.

Radford 2001: Bruce Radford, ICAP: A Make-Believe Market? Public Utilities Fortnightly, August 2001.

Sheffrin 2001: Anjali Sheffrin, Director of Market Analysis, Memo to the ISO Board of Governors, Market Analysis Report for June 2001, July 20, 2001.

Stavros 2001: Richard Stavros, Wall Street Rethinks Merchant Power, Public Utilities Fortnightly, Sept 15, 2001.

Sterman 2000: John Sterman, Business Dynamics, Irwin McGraw-Hill, 2000.

WA OTED 2001: State of Washington, Office of Trade and Economic Development, Energy Policy Division, Convergence: Natural Gas and Electricity in Washington, May 2001.

Lessons Learned from the California Power Crisis

A Discussion Paper for the California Energy Commission Workshop on Exploring Alternative Wholesale Electricity Market Structures for California, November 7, 2001

This paper contains the personal views of the author and they should not be regarded as the views of EPRI or its members. The purpose of this paper is to provide some background context for a companion paper "Comparison of a Competitive Wholesale Power Market with Alternative Structures through a Long Term Power Market Simulation Model."

Stephen T. Lee, Ph.D.
Area Manager, Grid Operations and Planning
EPRI
Palo Alto, California

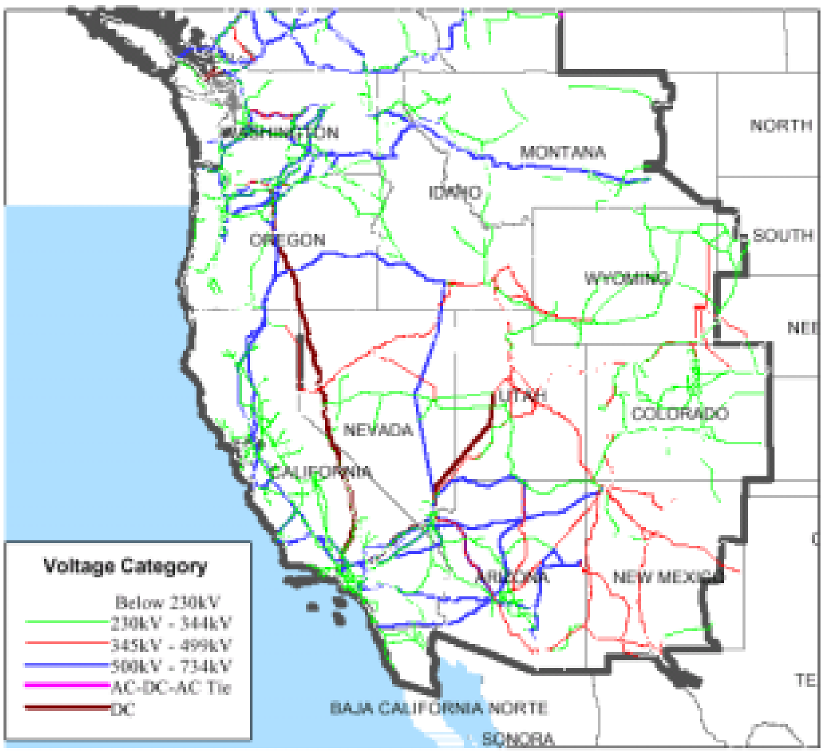
November 1, 2001

Table of Content

1. Background on California's Power Crisis

Background

The electric power system in California is part of the Western System Coordinating Council (WSCC), which is one of three Interconnections comprising the entire electric power system in North America, including Canada. The California power grid is primarily connected to the rest of WSCC through Oregon in the north, and Arizona in the east (and to a lesser degree with Nevada). California also has a limited tie to Mexico in the south. Figure 1 shows the position of California in the WSCC power grid.



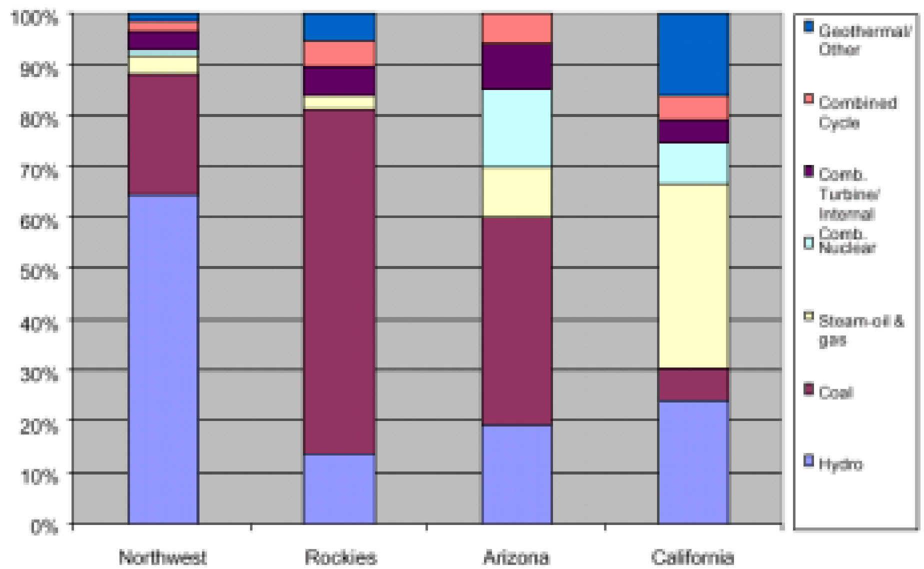
Source: Staff Report to the Federal Energy Regulatory Commission on Western Markets and the Causes of the Summer 2000 Price Abnormalities

Figure 1 - Western System Coordinating Council Power Grid

California, due to its long-held preference for less expensive hydro power in the Pacific Northwest and coal-fired energy east of its border, traditionally imports a large amount of energy from the north and from the east. Its transmission capacities from these two directions, although quite substantial, are limited.

California has three investor-owned utilities (IOU): Pacific Gas and Electric Company (PG&E) which serves northern California, Southern California Edison Company (SCE) which serves a large part of southern California, and San Diego Gas and Electric Company (SDGE) which serves the city of San Diego. In addition to these IOUs, there are municipal utilities, the larger of which include Los Angeles Department of Water and Power (LADWP) which serves the city of Los Angeles, and Sacramental Municipal Utility District (SMUD) which serves the city of Sacramento. These municipal utilities are not subject to the regulation of the CPUC and are not restructured.

The generation capacity mix in California, as seen in Figure 2, comprises about 25% of hydroelectric power, 45% of generation which burns natural gas or oil, 8% of nuclear power, and 6% of coal-fired generation, with the remainder coming from geothermal and other sources of electricity. Of special note is the high degree of dependence on natural gas and oil as a source of fuel.



Source: Staff Report to the Federal Energy Regulatory Commission on Western Markets and the Causes of the Summer 2000 Price Abnormalities

Figure 2 - Generation Capacity Mix in the Western System Coordinating Council Sub-regions

In comparison, the Pacific Northwest has about 65% of its generation in fairly inexpensive hydroelectric generation. However, its surplus of electricity for export to California depends on the amount of rainfall and snowfall, which may vary greatly from year to year. East of California, in Arizona and in the Rocky Mountain region, the amount of coal-fired generation ranges from 40% to 65%. The amount of export available from these areas would depend on their own demand for power and the random nature of power plants being forced out of service due to equipment failure when they are not on scheduled maintenance outages.

In addition to the transmission limitations for importing power into California, the internal transmission capacity linking northern and southern California is also limited. Known as Path 15, the limitation here could restrict the amount of power that can be sent from southern California to northern California even if surplus generation from the south is available to help meet the demand in northern California. This was most apparent on January 17 and 18, 2000 when northern California had to suffer from rotating blackouts when generation was available in the south.

History of Deregulation in California

In the 1970s, the global oil crisis caused a sharp increase in the prices of electricity in the U.S. It led the U.S. Congress to pass a law in 1978, the Public Utility Regulatory Policy Act (PURPA), to require utilities to buy power from Qualifying Facilities, with emphasis on encouraging renewable energy and cogeneration, at prices based on Avoided Costs. The California Public Utility Commission (CPUC) approved contract prices for these sources of electricity based on projected future oil prices. The consequence was that California utilities were saddled with high prices for these contracts, which were passed on to customers.

The Energy Policy Act, passed by Congress in 1992, created a new class of generators to compete with the utilities and mandated the Federal Energy Regulatory Commission (FERC) to open transmission access for the independent power producers. Subsequently, in 1997, FERC Orders 888 and 889 required utilities to provide open access and to post non-discriminatory transmission tariffs through the OASIS (Open Access Same-time Information System) , which is accessed through the Internet. FERC also encouraged the formation of Independent System Operators (ISO) to maintain a fair and reliable operation of the regional power grids. By late 1999, almost two dozen states have passed restructuring legislation.

Coming back to California, it was among the first state to embrace deregulation. Compounding the effect of oil prices and contract prices for Qualifying Facilities, the electricity rates in California were exacerbated by the huge cost overruns in the construction of nuclear power plants during the 1980s. This caused the large electricity users to consider expanding or moving their operations outside of California and created a political pressure to lower electricity rates. In 1995, the CPUC set out to restructure the investor-owned utilities (IOU) in California. In 1996, the State Legislature in a period of a few months, passed the Assembly Bill 1890 to deregulate them, with the intention to introduce supply competition from outside the state and to break the monopoly of the IOUs.

As part of the deregulation terms, consumers were given a 10% cut in electricity prices and a rate freeze until March 2002 or until the utilities have recovered all their past investments from a transition charge, whichever happens first. The utilities were allowed to issue $7 billion in bonds to pay for the 10% rate cut and charge customers over 10 years to repay the bonds.

To break the monopoly of the utilities, they were to sell half of their power plants, turn grid operation to an independent system operator (the California ISO), and retain the distribution business.

Nobody at the time could see the looming problems of power shortage a few years afterwards. Power supply was more than adequate in California at that time, and demand growth was low. Deregulation appeared to be the best way to reduce electricity prices in California.

The California Power Market Structure

Before restructuring, the California power industry was vertically-integrated and regulated by the California PUC. Vertical integration means that a single company owns and operates the entire system of generation, transmission and distribution. It had the responsibility to optimize the entire system to provide electricity to customers reliably and economically. It also had the ultimate obligation to serve the customers and avoid blackouts according to industry-accepted standards. Although it had a monopoly for supplying electricity in its service area, it was subject to a system of checks and balances, with the public utility commission scrutinizing every rate-increase application and the non-governmental consumer advocates challenging rate hikes in an open forum. The PUC also had the power to regulate the utilities financially through rate adjustments for imprudent management decisions, e.g., in fuel contracts, power purchase contracts, power plant construction cost overruns, etc. The utility was expected to promote demand-side management and conservation, hedge against fuel supply uncertainties, and pursue least-cost integrated resource plans. As a compensation, the utility's financial risk was reduced by a guaranteed reasonable rate of return on its investment and a quick pass-through to the consumers of fuel cost increases.

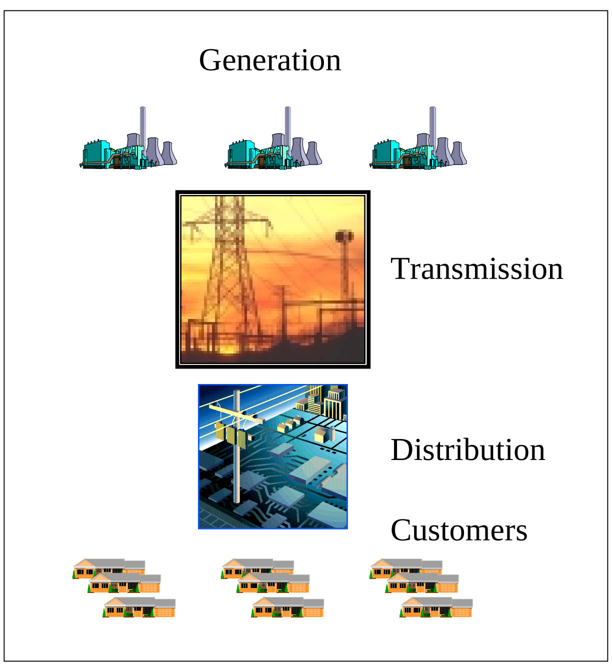


Figure 3 - Vertically-Integrated Utility Company Before Deregulation

Figure 3 illustrates the structure before deregulation.

There were criticisms about the regulated structure. Without competition, the utilities were not motivated enough to minimize their operating costs for maximum efficiency. Cost overruns on construction projects were a problem and fear of imprudence charges might have delayed the cancellation of projects when demand growth started to slow down. Thus, they faced charges of goldplating and over-investing. The surplus in generation capacity in the United States during the early 1990s was the result, although that surplus was conducive for the later market competition. In retrospect, with recent events, reasonable over-investment does not seem so bad.

The California deregulation bill resulted in a market structure that is very different. It is illustrated in Figure 4.

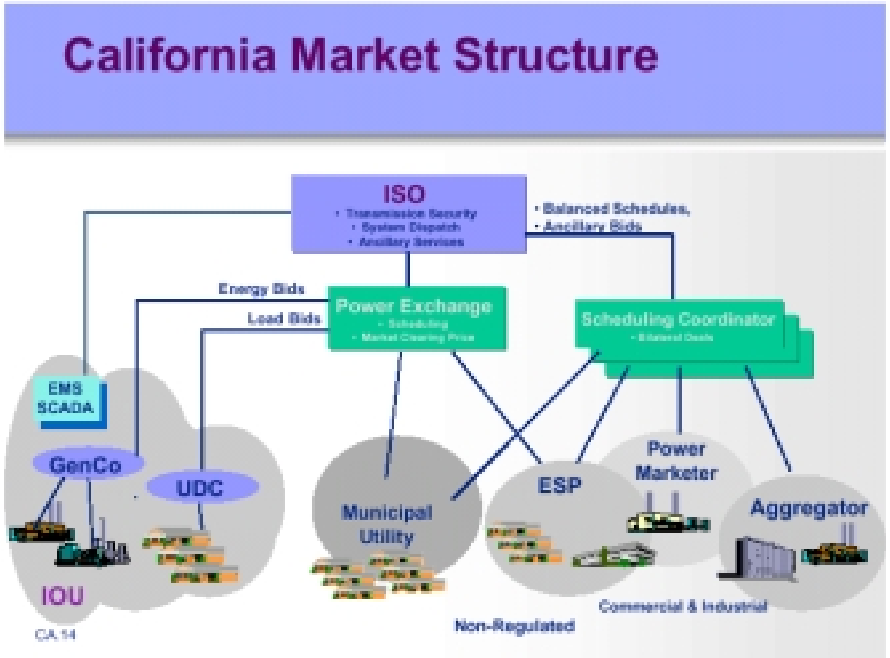


Figure 4 - California Power Market Structure After Deregulation

This market structure was designed to let the competitive market provide electricity freely, without interference from the State. The law broke the monopolistic IOUs into generation companies, and transmission and distribution companies. To create competition on the supply side, the IOUs were forced to buy all of their electricity from the Power Exchange (PX) and to sell all of the electricity produced by their remaining power plants into the Power Exchange and they were not allowed to enter into long term power purchase agreements.

In this new structure, the California ISO is charged with the responsibility to maintain reliable operation of the transmission grid by matching the system demand with the market supply, according to schedules provided by the Power Exchange. It also operates an ancillary services market to provide operating reserve in the power grid to maintain reliability.

The function of the Power Exchange is to determine the market clearing prices, in the Day Ahead Market, Hour Ahead Market and the Real Time Market.

Figure 5 shows how the California power market was designed to work. There would be certain amounts of bilateral trades with outside regions. There would be block forward energy purchases that are essentially longer term purchases through either the PX or a private exchange called the APX. A large part of the electricity would be cleared through the PX a day ahead and an hour ahead through a bidding system. Only a small part of the electricity (less than 5%) would be scheduled by the ISO in the real time to balance the energy market, due to uncertainties in demand forecasting.

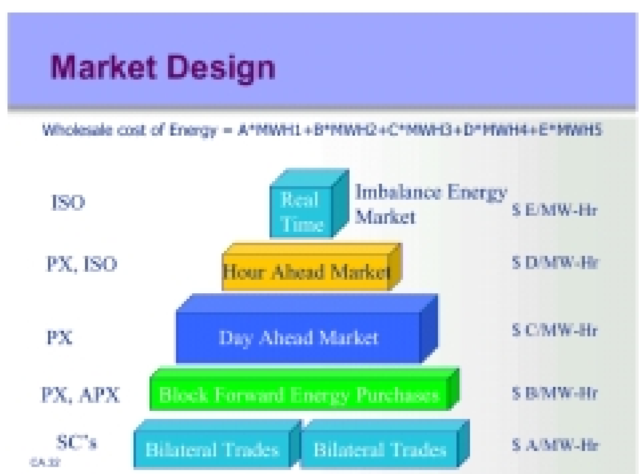


Figure 5 - California Market Design

The way market prices are determined is through matching demand bids with energy offers within the same market. The marginal price at which demand matches supply then sets the market clearing price, which is applied to all suppliers equally. This single market clearing price is a theoretically most efficient pricing signal for a competitive market. It provides a financial incentive for the market to recoup its past investments and to invest in new power plants for future needs.

As a safeguard against extraordinary price spikes, the

ISO has the power to set a price cap (upper limit) on the market clearing prices. When normal energy prices were in the range of $20-30 per MWh, the price cap of $750 per MWh seemed to be out of the realm of possibility.

Now that the original California power market structure has collapsed, many good questions are being raised. Was this market system complete? Where were its design flaws? Did it provide all the checks and balances that were found in the vertically-integrated system before deregulation? Will the competitive market voluntarily provide for capacity needed in the future? What kind of market structure will provide for a rational and economical expansion of transmission capacity? How would it provide for social benefits and adequate public safety? How would it provide for ancillary services? Do policy makers need to rethink the approach to deregulation? What role can a public power authority play to complement a competitive power market? The list of questions go on and on.

2. Analysis of California's Power Crisis

EPRI held a workshop on the Western States Power Crisis in June 2001 and published a white paper on the analysis of the problems and the proposed solutions [1]. The key subjects that relate to this paper are briefly treated below.

Obligation to Serve and Capacity Shortage

In retrospect, the power crisis of California is a predictable consequence of supply and demand imbalance. During the early 1990s, demand growth in California was very slow, and California had excess generation capacity. In the last few years, the economy in California has been booming, resulting in a surprising strong growth in demand for electricity, especially due to the fast penetration of high-tech companies. However, even though the capacity tightness was known by the CA ISO as early as 1999, the deregulated market structure in California did not give the ultimate responsibility and authority to a single entity to maintain supply and demand balance, except in the real-time operation realm to the CA ISO. In other words, the market was assumed to voluntarily build enough generation capacity to meet future demand of electricity because of profit incentives. No provision was made in the event the market fails to provide.

In the regulated power industry, the utilities have the obligation to serve. After deregulation, the transmission and distribution companies are not obligated to build new power plants. The CA ISO was seen as the power procurer of last resort, but mainly in the daily operation. Therefore, when the market failed to build enough power plants to meet a particular state or region's electricity demand, there was no organization ready to take action to ensure proper supply and demand balance for a well-functioning competitive market.

Electricity is not a commodity like wheat or water. It is not quite like air travel or telephone service. Electricity cannot be easily or efficiently stored in large quantities by consumers and producers, unlike wheat or water. Electricity provides essential services, like heating, lighting, cooling, business operation, medical health and public safety. It cannot be easily substituted with other energy sources at the consumer level. Not like air travel. If air fare is too high, a traveler can defer the trip or take a train instead. The differences between electricity and other services should be recognized when designing a competitive market. Price volatility is a known byproduct of competition, especially at times of large supply and demand imbalances. The consequences of price volatility on consumers are, however, different depending on the commodity or the service in question.

Other features of electricity that are different from other commodities or services are the long lead time and high financial risk associated with building new power plants and new transmission lines. Smaller power plants that burn natural gas or fuel oil, especially combustion turbines, can be built in about two years. Larger power plants that burn coal or use the combined cycle design take three to five years. Added to the construction time is the lead time required for obtaining all the licenses and environmental permits before construction is allowed. Such lead time often took two or more years in California. However, by obtaining permits ahead of time and banking them, a generating company can reduce the construction decision lead time to a minimum of two to three years for gas-fired units.

Market Power

The question whether market power was exercised during the power crisis in California was a subject raised by various researchers, e.g., by Joskow and Kahn [2]. Investigation was made by FERC into this possibility and the conclusion was that no evidence could be found to support this allegation. However, skepticism persists among the general public about market manipulation. Possible actions which people felt that generators could have taken to manipulate the market include withholding power from the Power Exchange market, putting generating units on non-operational status, and charging excessive prices.

The study by Joskow and Kahn attempted to identify the price gap between reasonable marginal costs of producing electricity and the actual market prices. However, significant price gaps are not illegal in a competitive market if they are fair and reasonable.

For this paper, it is important to note that regardless of the legality of market prices, in a scarcity market where the demand elasticity is low, the law of supply and demand will result in very high prices. The demand elasticity of electricity is basically low unless a real-time pricing scheme is implemented which allows consumers to know the real-time price and to get a lower electric bill if they shift their use of electricity to the lower-price hours. In the case of California, not only was real-time price not available, the electric rate was actually frozen during this whole period of wholesale price increases.

In a scarcity market, if price control is imposed, it will result in shortages and would require rationing. Sustained price control will prolong shortages if the price cap does not allow sufficient profit incentive for creating new supplies.

Financial Complications

The California power crisis also brought out the financial problems when deregulation went wrong. The rate freeze during the transition period had the unanticipated effect of causing a huge financial burden on the three IOUs when the wholesale energy prices went up dramatically since June 2000. Because the deregulation law requires them to sell all of their generation into the market and buy all their electricity from the market, and they were not allowed to hedge against the market prices by entering into long term power purchase agreements, the distribution companies of the three IOUs were completely exposed to price volatility. However, the increases in the purchase costs of PG&E and SCE could not be passed through to their customers because of the rate freeze. As a result, they went into an undercollection situation totaling about $12 billion.

The financial picture is somewhat more complicated. The generation company part of the two IOUs made money by selling their electricity to the market. Because of restructuring, the generating companies are legal entities which are separate from the distribution companies, although they are owned by the same parent companies. In effect, one side of the parent company is losing huge amount of money while another side is making money.

When PG&E and SCE did not have enough cash to pay their debts, their credit ratings were dropped to the level where they could not purchase electricity on credit. It became necessary for the state to start taking over the cost of buying electricity for them to distribute. For some time since taking over, the state had been spending about $40 million each day to buy electricity. Only when the state had negotiated enough long term power purchase contracts and the California electricity consumers had responded rigorously to the appeal for conservation, did the wholesale prices come down to a normal level.

Another complication in the financial matters is that any idea about reversing deregulation would be very expensive. When the IOUs sold half of their power plants, the market prices they received were much higher than the book value of those depreciated assets, with prices that range from 2 to 4 times book values. Therefore, if the state wanted the utilities or a state agency to buy those power plants back, the buyback prices would be very high. The consequence would be that the state will have to pay again for those same power plants which consumers had previously paid fully or partially already.

Solutions Considered in California

The solutions or options that are considered or suggested for solving the California power crisis are many. They can be generally classified as follows:

- **Solve supply and demand imbalance**
  - build new generation
  - sign long term power purchase contracts
  - conservation
  - demand management including price responsiveness
  - build new transmission
- **Solve financial problem**
  - raise electricity rates
  - rescue the distribution companies from bankruptcy
  - State buys power for the distribution companies
- **Change the market rules (suggested by FERC)**
  - Eliminate requirement that the three IOUs must sell and buy all their power needs from the PX.
  - Require market participants to schedule 95% of transactions in the day-ahead market. Propose a FERC penalty for deviations above 5% of the hourly load requirements.
  - Allow single price auction (i.e., the marginal price applies to all cleared market transactions) only at or below $150/MWh. Above it, accept individual bids if needed to clear the market, require reporting to FERC with cost information and subject to refund, if the price is found to exceed opportunity costs
- **Restructure the power business in California**
  - State forms a power authority to build power plants, as in other states such as New York.
  - Cities form municipal utilities like in Los Angeles and Sacramento.

In the rest of this paper, the author will describe a computer model of the California power market which has been exercised to provide some understanding and insights about the potential effects of some of these suggested solutions.

3. Market Competition and Restructuring

Objectives of Market Competition and Restructuring

Having looked at the power crisis in California, it is useful to re-examine the theory of market competition and restructuring and to consider what lessons can be learned from the California experience. First it would be appropriate to re-examine the objectives of market competition and restructuring in the case of electricity. Among them, there are possibly the following three objectives:

1. Lower electricity cost and improve efficiency

   This is the most often quoted objective of market competition and deregulation. The question is whether empirical data or simulation models support these claims. Another question is whether market competition is the best way to lower electricity cost and improve efficiency. Before deregulation, utilities took a comprehensive integrated planning approach to optimize their power system, from fuel supply, power plant mix and unit sizes, transmission and distribution investment, all the way to distributed generation and demand-side management. For researchers, the effectiveness of market competition in comparison with the command and control functions of a regulated vertical utility structure has yet to be studied with sufficient data and computer modeling.

2. Encourage innovation in technology by profit incentives

   In order to lower electricity cost in the long run, it is important to make use of innovative technologies that enable lower costs. Market competition has been proven in many fields to be the most effective means of encouraging innovation. However, it should not be overlooked that long-term R&D is needed to develop innovation, requiring commitment and major resources, which individual enterprises may not afford to do.

3. Encourage capital investment with potential higher profits

   A competitive market depends on profit potential. If the profit potential is not high enough to compensate for the risk of investment, the free market will avoid such investments. However, in the long run, providing higher potential profits to the market contradicts with lowering electricity cost, unless new and more efficient technologies are introduced. In a regulated industry, all savings due to new technologies get passed on to the ratepayers. In a competitive market, investors and consumers share these cost savings. When a market is dominated by a few large players, consumers may lose out. This means some regulation mechanism is needed to ensure sufficient competition.

Potential Risks of Market Restructuring

The power crisis in California brought out some potential risks of market restructuring in the electricity industry. In re-examining market restructuring, a good question to ask is, "Does market restructuring and the ensuing competition introduce new or greater risks into the power industry?" It seems that the following risks are now in evidence if a healthy and robust power market structure is absent.

1. Real-time or actual electricity price may become more volatile.
2. Blackouts may be more frequent.
3. Social impacts may result from price spikes and blackouts.
4. Power companies may go bankrupt.
5. Government and taxpayers may have to bail out financially-troubled power companies.
6. The economy may go into a recession.

7. The deregulated power market may be prone to an unstable business cycle in the absence of a complementary regulating policy.

An analysis of the electricity prices in California over the last 25 years seems to say that historical electricity price has become more volatile after deregulation. Figure 6 shows that during the oil crisis in the 1970s, when oil prices went up by a factor of three, electricity prices in California went up by a factor of two. This happened over seven years, from 1975 to 1982.

Because the current situation in California has not settled down, the eventual electricity rate increases to customers are not yet known. They will depend on how much of the IOUs' undercollection, and the state's power purchase cost will be paid by the customers, how much will be paid by the taxpayers (through the state), and how much will be financed with bonds to be repaid by customers over a period of time. However, we can make an estimation based on the wholesale energy costs at the actual wholesale energy prices reported by the CA ISO since summer 2000. The curves for electricity prices for 2000 and 2001 in Figure 6 were therefore estimated values only.

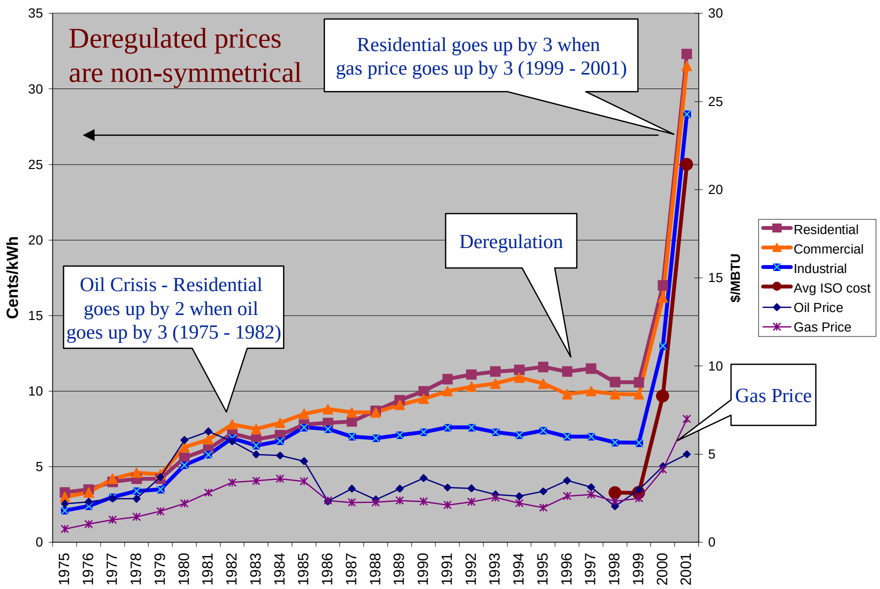


Figure 6 - Electricity Prices in California from 1975 to 2000

Given those estimated values, one can observe that, after deregulation, assuming that the recent wholesale energy prices are passed through to the customers eventually, the residential rate at the beginning of 2001 would be about 32 cents/kWh (read from the left axis), as compared to 10.6 cents/kWh in 1999, in other words, by a factor of three (3), in less than two years. Over this same period, the average natural gas price went up by a factor of three.

It appears then that electricity prices in California went up higher and faster in this latest power crisis than during the oil crisis. This indicates greater price volatility.

On the exposure to boom-bust cycles, Andy Ford [3] and this author have recently done work with computer modeling that shows the propensity of a competitive power market to exhibit this phenomenon. The results of these simulations are part of the subject matter for this workshop.

The potential risk of boom-bust cycles to the industry is the possibility of generation companies finding themselves invested in power plants in an overbuilt region and not be able to compete in the surplus market. Financial failures of the weaker generation companies may result.

In a competitive market, generation companies have a financial objective to build the lowest cost power plant designs in the shortest lead-time in order to maximize profit. This is of course the fundamental principle of a free market, from which much prosperity have resulted for the world economy. When the economics are very clear cut in favor of one type of power plant, there is a higher risk of exposing consumers to dependence on a single fuel, than before deregulation. This is what has happened in recent years. Most new power plants built in the last few years burn natural gas. The natural gas prices in the winter of 2000 dramatically show how dependent on natural gas the power

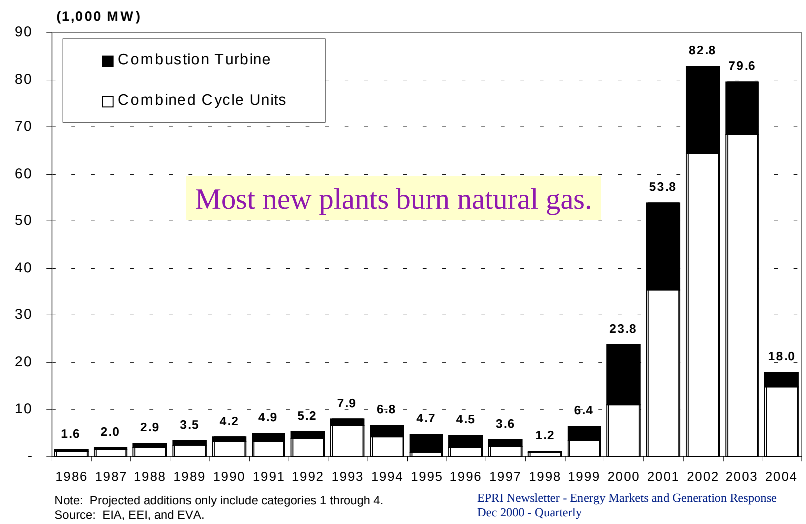


Figure 7 – Historical and Projected Combustion Turbine and Combined Cycle Capacity Additions

producers have become. This trend is expected to continue for at least a few more years, as can be seen by the types of new power plants that are under construction. See Figure 7.

Finally, a last question represents a hypothesis by the author which deserves academic research. The question is, "Can a public power agency with a limited role in the power market provide a stabilizing factor to dampen the potentially unstable business cycle and what is that role?" Another question is, "What type of action by a public power agency may cause financial disincentive to the private generating companies to build new capacity?"

Impediments to a Healthy and Robust Electricity Market

Continuing the re-examination of market competition, one should ask the question, "What are the impediments to a healthy and robust electricity market?" The following points deserve to be considered.

1. Electricity is a necessity -- for light, heat, business and public safety.
2. Electricity cannot be substituted, e.g., train for airplane.
3. It takes a long lead time to build power plants & transmission lines.
4. The entry cost & risk for new competitors are high.
5. Market imbalance takes time to change.

Policies, market structure, and market rules can be designed to overcome or reduce the impact of these factors. For example, social programs can be coordinated with the electricity business to ensure essential services to needy people when electricity prices are high. Public safety measures can be improved and coordinated with electricity emergency operation to minimize social and public disruptions during outages. Electricity may be substituted by other forms of energy sources, if alternative technologies are available, e.g., distributed generation.

Demand-side management and real-time electricity pricing would increase demand elasticity and enable customers, both large and small, to respond to the market conditions, thus minimizing their electric bills. Industry and commercial customers may adjust their work schedules and operations to shift more of their electric usage to hours of lower market prices. Residential customers can also change their dish-washing and clothes-washing times to take advantage of time-of-day pricing.

Regulatory agencies may streamline the permitting process to shorten the lead time to build power plants and transmission lines. The entry cost and risk for new competitors may be reduced by developing alternative energy sources and electricity delivery mechanisms, such as direct-current micro grid. Financial risk may also be reduced by improving the way fixed costs can be recovered by the power producers.

Finally, a government agency or an independent agency may be empowered to monitor regional balances of electricity supply and demand, and pursue timely options to maintain the long term balance of electricity supply and diversity of fuel dependence, for public good.

If these measures are taken in a comprehensive manner, they should reduce the likelihood of design flaws and potential crises in the electricity market.

Can Market Competition Be Complemented by Government Actions?

Among the lessons learned from the California power crisis is the possibility that market competition could be complemented by government actions.

In addressing the power crisis, the California state government did take some actions. The question is what types of government actions would be complementary to the market.

In reality, no competitive market is without some form of regulation. The term regulation should be used in a broader sense than price control, or the establishment of regulated monopolies, in the case of the electricity market. Just like in the monetary market where the Federal Reserve Bank regulates money supply and adjusts interest rate to guide and control the economy, in an electricity market, it is conceivable that the government could implement some form of action that guides and supplement the regional or national power market.

What these forms of action should be is an area for discussion and research.

Foremost, in my personal opinion, is the need to maintain a balanced regional supply of capacity, both in generation and in transmission. The market may not be relied upon entirely to provide sufficient reserve margins. The government or an independent and empowered entity should consider monitoring the market and taking actions to build new plants or incent them to be built when the market fails to provide, for whatever reason. Because the consequence of power shortage is great, and the lead time to correct the imbalance in an economic manner is long, it may be necessary for a carefully designed plan to prevent electricity shortages.

Another is the need to pursue public benefits. Private enterprises do not ensure public benefits as their primary goal. Government plays a major role. Fuel diversity may not be achievable in a purely competitive market. In order to minimize dependence on a particular fuel, the government should consider taking actions to encourage fuel diversity. It can either provide financial incentives to the market to build different types of power plants or build them itself.

It is not the purpose of this paper to present a thorough discussion of this subject. What is hoped to be achieved is that this subject be considered and discussed among researchers and policy makers, so that a well-designed structure can be developed for implementing minimum but effective government actions to guide and assure the supply balance and stability of the power market. The companion paper [4] presents preliminary results from a long-term power market simulation model which may shed some light on these questions.

References

1. EPRI, “The Western States Power Crisis – Imperatives and Opportunities,” an EPRI White Paper, June 25, 2001.
2. Paul Joskow and Edward Kahn, "A Quantitative Analysis of Pricing Behavior In California's Wholesale Electricity Market During Summer 2000", November 21, 2000.
3. Andy Ford, “Waiting for the Boom: A Simulation Study of Power Plant Construction in California” (to appear in Energy Policy, Vol. 29, p. 847-869), March 2001.
4. Stephen Lee, “Comparison of a Competitive Wholesale Power Market with Alternative Structures through a Long Term Power Market Simulation Model,” November 1, 2001.

Comparison of a Competitive Wholesale Power Market

with Alternative Structures through a

Long Term Power Market Simulation Model

A Working Paper for the California Energy Commission Workshop on Exploring Alternative Wholesale Electricity Market Structures for California, November 7, 2001

Stephen T. Lee, Ph.D.
Area Manager, Grid Operations and Planning
EPRI
Palo Alto, California

November 1, 2001

Table of Content

1. The EPRI Long Term Power Market Simulation Model

This is a paper reporting on *work in progress*. It is prepared for discussion purposes at the November 7, 2001 California Energy Commission Workshop on "Exploring Alternative Wholesale Electricity Market Structures for California." A companion paper on "Lessons Learned from the California Power Crisis" is intended to provide some background context for this discussion [1].

Objectives of the Long-term Power Market Simulation Model

There are four objectives of developing and applying the EPRI Long Term Power Market Simulation Model (LTPMS):

1. Improve understanding of how a competitive power market works in the long term.
2. Study the potential price volatility and boom-bust cycles of the power market.
3. Study alternative market designs and the role of a Power Authority to complement the market.
4. Compare the long-term impacts of various market structures on the end-users and society.

EPRI does not provide policy advice nor advocate public policies, but rather results of objective research. It is therefore imperative that the assumptions and limitations of this LTPMS model be understood by users of these results and also that sensitivity analyses be done and reported in this paper to show that conclusions may be highly dependent on the assumptions adopted.

In the course of preparing this paper, the author had consulted with Professor Andy Ford, who is the author of another long-term power market simulation model [2,3] and had compared the methodologies and assumptions of the two models.

Assumptions and Limitations of LTPMS

The LTPMS model is a prototype computer model. As such, it is still undergoing improvements. Results of this model, however, have exhibited reasonable phenomena and quantitative data that can explain recent historical experiences in California. The results are corroborated in various degrees by the simulation model developed by Andy Ford, even though the two models use different modeling methods and look at somewhat different system data and assumptions. As preliminary studies, the results reported in this paper should be viewed as providing potentially useful insights, but not necessarily as definitive. This caveat especially applies to the numerical values of the data reported.

The modeling assumptions are now described.

General Assumptions and Methodology

The LTPMS uses the Microsoft Excel software in the form of many integrated worksheets in two different work book. One work book models a purely competitive power market. A second work book models a hybrid power market. The simulation uses a monthly time step over a 20 year period (240 months).

The following entities are modeled in separate worksheets:

- Grid Operator
- Independent Power Producers (IPP)

- Electricity Customers
- Power Authority (in the hybrid power market model)

There is also a worksheet for general data, and a worksheet for the economic parameters for new generation. Many charts are set up that display the input data and the results graphically. The feature of Excel which allows parametric tables to be set up for running a set of sensitivity analyses is extensively used, and the results of these sensitivity analyses are displayed in charts.

Grid Operator and System Model

The electricity system modeled in LTPMS is based on the California generation mix (the entire state), and historical load. Monthly load is represented by the maximum MW demand and the month's total MWh energy. It models the generating capacities in the categories of hydro, nuclear, coal, others, gas, oil and CT and not as individual generating units. It is not intended to be a detailed production costing model. Rather, its main purpose is to estimate a monthly average market clearing price for wholesale electricity. It does it by derating of capacity to account for maintenance outages and average forced outages. Hydro capacity is modeled by average hydro conditions and as derated monthly capacities. A simple load duration curve approach is applied for estimating the monthly capacity factors of the various capacity mixes. The assumption is that the competitive market generally results in producing electricity from these different capacity mixes in a utilization order according to their average operating costs.

The market structure is modeled after the California market before the California Power Exchange ended operation. There is a grid operator representing the CA ISO. The IPP worksheet in LTPMS represents the entire supply side of the California generation mix. Based on the load and available generating capacity for the month, the marginal energy cost of the entire system, marked up by 5% (user input), is assumed to be the average market clearing price for the month, if the reserve margin for the month is outside the zone of market power. The market clearing price is then applied to all sellers into the market for pricing the cost to the consumers.

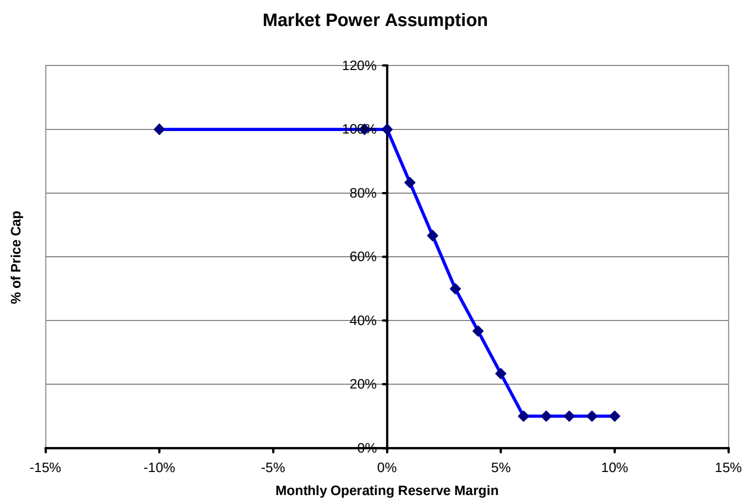
The percentage of the system load subject to the market price is an input data which can be changed. That is used to represent the availability of bilateral contracts. For example, 85% of the entire system load may be modeled as under bilateral contracts, with an average contract energy rate of $65/MWh. The terms of the contracts is over 10 years, with a second 10-year renewal term. During each term, the energy rate is fixed. For the second contract term, the energy rate is assumed to be adjusted for a 3% annual escalation for ten years ($87/MWh in the base case). It is further assumed that the bilateral contracts are supplied from the same capacity mix as in the entire system, in their relative proportions. This means that the market clearing price continues to be determined by the same capacity mix. However, the effect of the bilateral contracts is to provide a steady and fixed source of energy at a pre-determined price to the customers, thus reducing the effect of the market exchange price volatility on the entire energy demand of the state. It also provides a steady stream of revenue to the IPPs.

In the event that the market supply does not meet the customers' demand for electricity, the market shortfall is assumed to be met by the Grid Operator through out of the market purchases at the price cap price. For the base case, the price cap is assumed to be $250/MWh escalating with inflation at 3% per year.

Independent Power Producers

The assumption about the bidding behavior of the independent power producers has already been mentioned, i.e., under normal reserve margin conditions, they bid their marginal energy costs plus 5%. Based on the macro relationship deduced from the California experience and from academic short-term market simulator games conducted in Cornell University with students [4], it is felt to be reasonable to postulate a functional relationship between the available reserve margin for a month and a market-power index. In the work done at Cornell University, it was found that after a number of market simulations, the students would find out that they would make more profits by withholding power or bidding at the prevailing market clearing price, and that when the competitive degree is low, all bid prices rise to the level of the price cap. This is the case whether the market pricing is "as-bid" or "uniform."

In the LTPMS model, the market power is expressed as a percentage of the price cap. The assumed curve is shown below. The interpretation is that if the reserve margin is negative, the market power is assumed to be 100% of the price cap. Between 0% and 3%, it drops from 75% to 50%. From 3% to 6%, it goes from 50% to 10%. Therefore, depending on the reserve margin for the month, if the market power curve shows a higher price than the normal bid price of the IPPs (marginal energy cost plus 5%), that higher price will be assumed to be the market clearing price. The values of this curve may be changed by the users.



It is important to distinguish between this type of macro modeling of empirical relationships from any implication that this is making the assumption that IPPs set their bidding prices based on reserve margin. It is simple a reflection on the behavior of the market clearing mechanism of any competitive market when the supply curve and the demand curve comes together during periods of scarcity. It is also a reflection on the financial realities of IPPs having to recover sufficient revenue to permit past investments to be recouped and for new investments to be funded.

The LTPMS does not model any capacity payments to the IPP for new generating capacity. It does not have a ICAP market structure, like it exists in some power markets. It is a feature that would be useful to add to the LTPMS.

The LTPMS is different from typical production costing models in that the decision process of the IPPs in building new plants is directly modeled. It is based on a benefit cost calculation of a new gas-fired generating unit. The benefit is the current wholesale market clearing price minus the marginal energy cost of the gas-fired generation, expressed in $/MWh. The cost is the levelized fixed charges of the new gas-fired unit, expressed in $/MWh by accounting for the current month's average capacity factor for gas-fired generation. If the B/C ratio is greater one, it means that operating income (revenue minus operating cost) from selling power from a gas-fired unit exceeds the levelized fixed cost of building the unit over 30 years, including a desired return on investment (ROI). The base case assumption for the ROI is 15%. With the assumed life of 30 years for the power plant, and an assumption of 3% per year for other fixed costs, such as property tax and fixed operating costs, the base case levelized fixed charge rate is 18.23%. It is further assumed that the IPP as a whole puts 80% of the new capacity in combined cycle plants and 20% in gas-fired combustion turbines. These percentages may be changed in the model.

The LTPMS computes the B/C ratio each month. It assumes that the IPPs are of two types. One type already owns the generating capacity in California. The other type represents new entries in the California market or project-based entrepreneurs. The first group is assumed to have an interest to see that the power supply in the market maintains a reasonable reserve margin, so as to avoid market collapse. In the base case, this is assumed to be 12% reserve margin. Therefore, the decision of this group of IPPs involves the monitoring of the monthly B/C ratio and the estimation of how much new generation capacity should be built so that in 3 years (the construction lead time), the new capacity plus construction already in progress will meet the future peak load forecast with the desired reserve margin. If the current month's B/C ratio is less than one, there is no decision to build. If it is greater than one, it will build only to the amount not exceeding the target reserve margin, as described. Because of the seasonal load patterns, the monthly B/C ratio is highest during the summer months and the average reserve margin is also lowest during those months. It is assumed that the decision to build is made in the summer and the new plants will come on line in May of the third year.

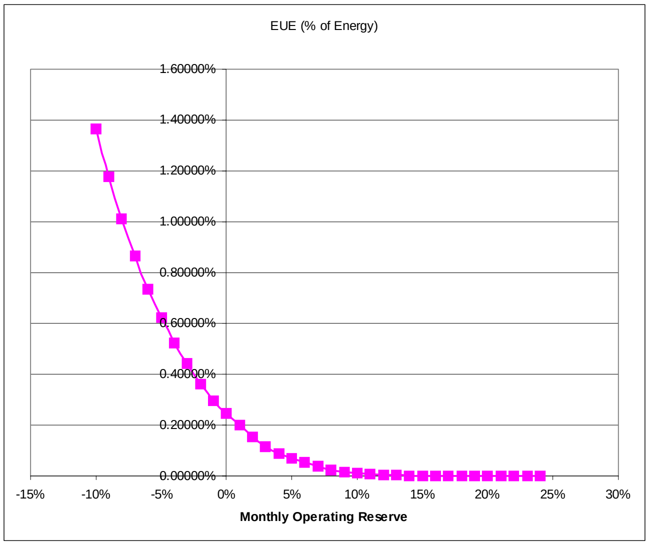


Figure 8 – Expected Unserved Energy vs. Monthly Operating Reserve

As for the second group of IPPs, it is assumed that they do not have the same longer term view of the first group. Instead their decision is to start building a certain number of gas-fired units every summer when the B/C ratio is greater than one. An input parameter to the LTPMS model is the number of these units that will be built. This parameter is called the Competitive Degree. The sensitivity of the simulation results to both parameters, the target reserve margin and the Competitive Degree, is tested and reported in this paper. These are key assumptions for the model. As a model of human decision making, it is necessarily not precise and is subject to disagreement among modelers. However, by parameterizing this decision process into these two input data, the LTPMS is able to show the sensitivity results and provide insights on what are reasonable assumptions for these parameters.

Electricity Customers

The LTPMS does not assume any retail price freeze. The wholesale electricity prices, from both the exchange market and the bilateral contracts, are assumed to pass through immediately to the electricity customers, on top of the fixed components of the rate structure. The average fixed rate is assumed to be 7 cents/kWh. It is assumed to phase out in March 2002 and then drop to 1.5 cents/kWh thereafter.

The customers are assumed to exhibit a price elasticity. The demand elasticity factor is an input value. For the base case, it is assumed to 0.2% for each % increase in the customer electric rate from 12 months ago. There are different ways one can model the customers' response to prices. The LTPMS model does not have hourly details to model RTP (real time pricing). It is felt to be reasonable to model a year-to-year price elasticity rather than month-to-month because of the potential volatility of monthly price fluctuations from month to month. The year-to-year relationship seems to replicate the effect of conservation seen in California during the summer of 2001. This model also indicates the potential of demand bouncing back in subsequent years when prices are lower.

To simulate the effect of blackouts on the customers and on the economy, a loss of load probability calculation for estimating the expected unserved energy (EUE) is set up in the LTPMS model by means of a lookup curve. This curve is computed by assuming that the entire capacity mix of the system consists of 100 generating units with an average forced outage rate of 8%. The actual load duration curve of the system in a summer month is then convolved mathematically with the generation outage probabilities to produce a curve of EUE as a function of the reserve margin of the month. This method, using actual unit sizes and forced outage rates, is routinely used in probabilistic production costing models. This curve is shown in Figure 8.

The cost to the society of these blackouts is assumed to be $100,000 per MWh of unserved energy. Given that price caps of $1000 per MWh is not uncommon in some power markets, it seems reasonable to assume that the cost to the society, including the effect of blackouts on the economy, may be as high as 100 times higher. A recent study by the Consortium for Electricity Infrastructure to Support a Digital Society (CEIDS) on the effect of blackouts and poor power qualtiy in California had shown impacts on its economy to be from $13 to 20 billion. Again, this being an important assumption, the LTPMS model performs sensitivity analyses on this input data.

Public Power Authority (PA)

The LTPMS model represents the public power authority (PA) in a the hybrid version of the model. The PA is assumed to have a role to build only peaking capacity, specifically, combustion turbines. Note that this is only one particular role that a PA can possibly play in a hybrid power market. The purpose of modeling the PA in the LTPMS model in this manner is to explore the potential interaction it has with the competitive market, and the effect on the customers and society. The basic premise in this application is therefore to see if the PA can reduce price volatility and enhance system reliability. Through the LTPMS model, it would be interesting to see if the PA's role will produce disincentives for the IPPs resulting in their reducing their building program. In other words, we wanted to study whether the PA can play a complementary role to the market and not find itself in a situation where it has to supply the entire market.

In the study, the PA is given the financial edge of being able to raise financing at 12% (3% lower than the 15% ROI for the IPPs). With 3% other fixed costs, the levelized fixed charge rate for a PA is therefore 15.41%.

In the hypothetical hybrid power market, the PA is assumed to act as a "contrarian." In other words, its decision to build or not to build combustion turbines is the reverse of the decision by an IPP to build or not. So, the rule is that if the summer B/C ratio, as computed by the IPPs, is less than one, i.e., when it is considered not profitable for the IPPs to build, the PA will start building or procuring combustion turbines. The limiting amount of the PA's building program is assumed to be a user-input percentage of the annual load growth, ranging from 0 to 120%. For example, a 60% value means that the PA will build an amount of CT equal to 60% of the annual load growth in MW from the previous year. A value of 120% would mean that the PA will essentially build for the entire state's annual load growth plus 20% reserve margin. The purpose of exploring the sensitivity of this input assumption is to identify whether there is an optimal share of the total generation capacity which complements the market and maximizes public good.

As for the cost recovery and operation of the PA's combustion turbines, it is assumed in the LTPMS model that the CTs run at a monthly capacity factor of 3% for 9 months, and 5% for the three summer months. The operating costs and the levelized fixed charges of the PA are assumed to pass through to the electricity customers at cost. The PA is not assumed to participate in the market bidding but rather operates according to a bilateral contract with the Grid Operator, at cost. The Grid Operator would combine these costs with the market costs and the total costs will be passed to the electricity customers in a total electricity rate.

Fuel Prices and Other Assumptions

The fuel prices in the LTPMS model can be changed by the user. The base assumptions are shown in Figure 9. It includes the natural gas price spike in the winter of 2000. Thereafter, the prices are assumed to increase at the annual escalation rates shown in Figure 9. Note that the feedback effect of high gas consumption on later gas prices is not assumed.

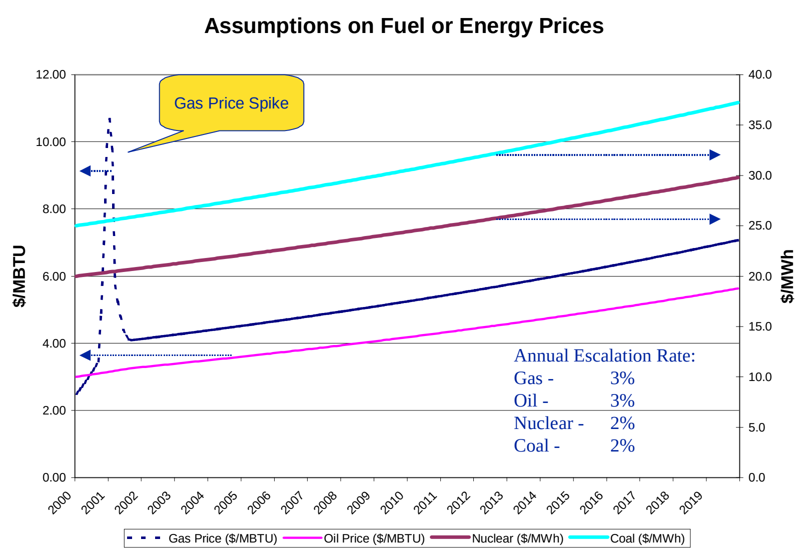


Figure 9 – Fuel Price Assumptions

The capitalized cost of a gas-fired combined cycle unit is assumed to be $800 per kW and that of a combustion turbine is $390 per kW.

Emission cost to society is also computed in the LTPMS model from the energy production of each generation mix. The NOX emission cost is assumed to be $30 per lb. It is noted that during the summer of 2000, the NOX emission credits were trading in the range of $5 to $40 per lb. The SO2 emission cost is assumed to $250 per ton, and the CO2 emission cost is assumed to be $1 per ton. These data are consistent with data reported by the U.S. Department of Energy.

2. General Results

The LTPMS model was exercised with a number of sensitivity runs around a base case scenario. The base case assumptions were noted earlier in this paper. The results of the model are shown in the following set of graphs. Figure 10 shows the 20-year capacity and demand picture along with the total costs to the public (electric bill and the social outage cost).

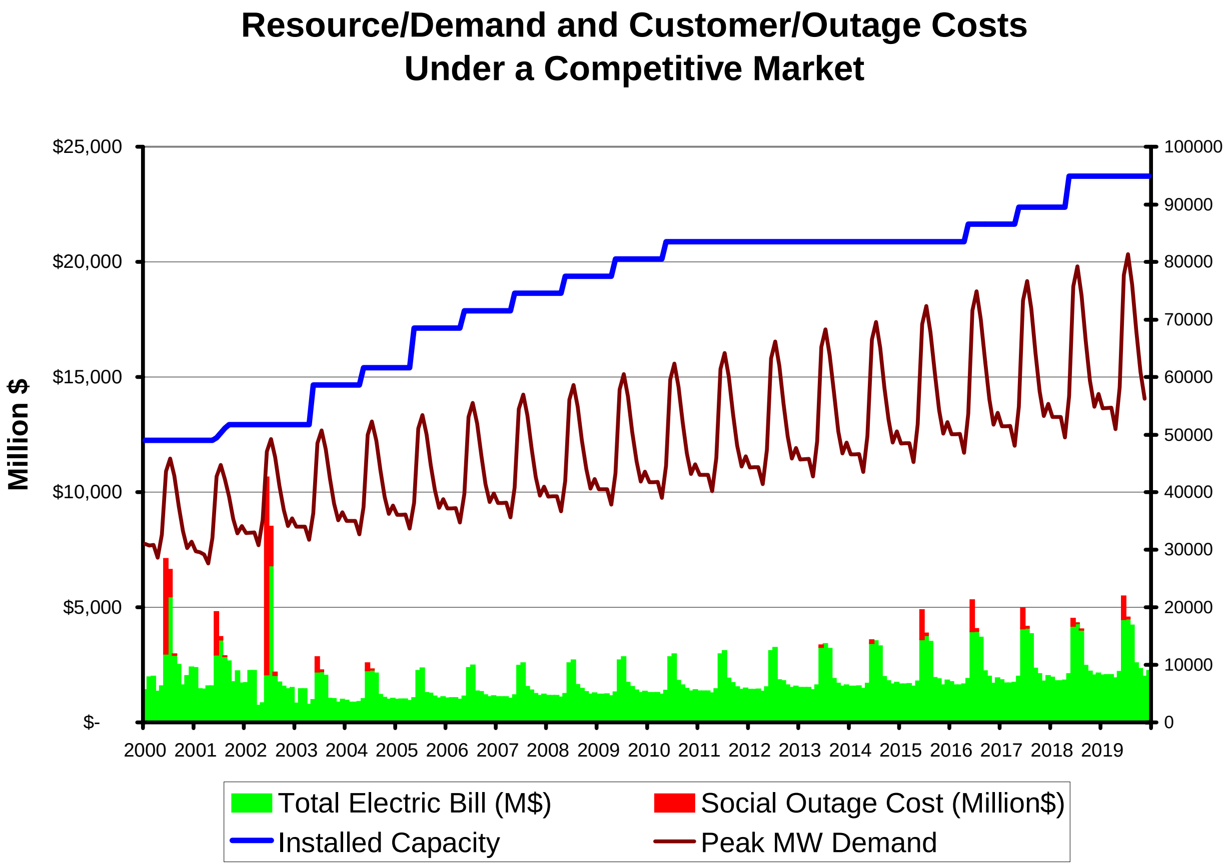


Figure 10 – Resource/Demand and Customer/Outage Costs Under a Competitive Market
No bilateral contract, demand elasticity = 0.2, competitive degree = 10, target reserve = 12%

Figure 10 clearly shows a boom-bust cycle. Following the capacity shortage condition of 2000 to 2005, a number of new units are put under construction and come on line from 2003 through 2010. Then a period of low market prices bring about a lull in construction until a new wave of new plants come on line from 2016 to 2018. It also shows that in the year or two just before a new wave of capacity, e.g., 2014-2015, the electricity rates are high, accompanied by significant social outage costs. This phenomenon seems to be related to the way the B/C ratio for new capacity varies as a function of the market prices, which increases when the reserve margin of the market drops to a low enough level. The competitive degree has an effect on the size of the boom-bust cycle. For example, Figure 11 shows the simulation results if the competitive degree is 2 instead of 10.

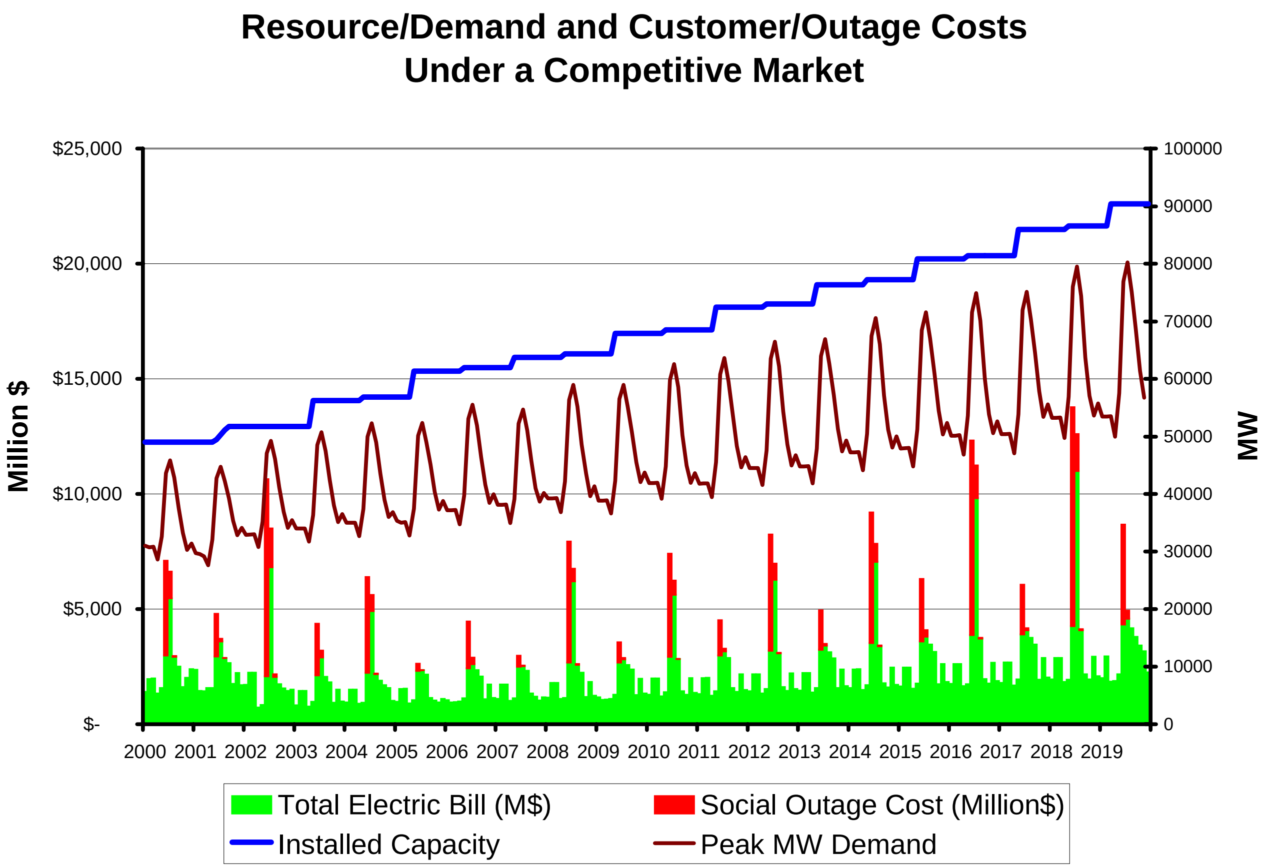


Figure 11 – Resource/Demand and Customer/Outage Costs Under a Competitive Market
No bilateral contract, demand elasticity = 0.1, competitive degree = 2, target reserve = 12%

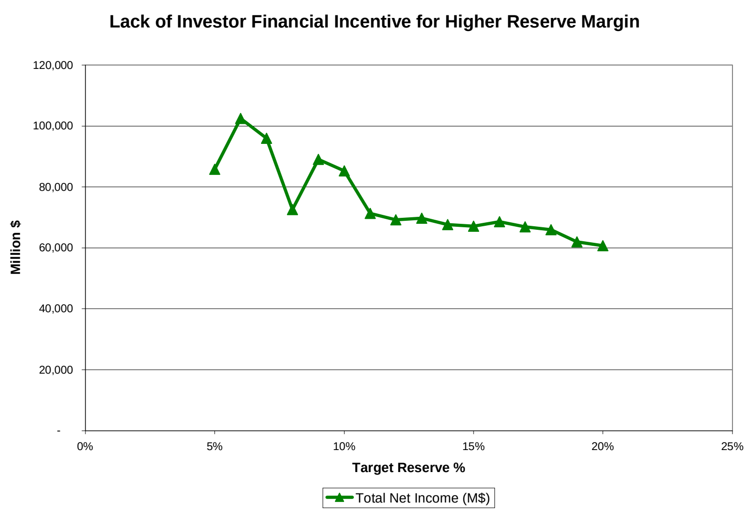


Figure 12 – Net Income for IPPs Over 20 Years vs. Target Reserve % Assumption

In Figure 11, the reduced level of competition makes the effect of the 12% target reserve margin very apparent. The boom-bust cycle has disappeared and what we observe is the steady building program of the IPPs towards a 12% target reserve in the entire market. This has the effect of keeping prices relatively high so that the financial incentive for building new capacity is maintained, while the social outage costs increase.

To explore what reasonable target reserve margins are to the IPPs, the effect of a number of these simulations, using different target reserve margin, on the net income over 20 years for the IPPs as a whole is shown in Figure 12. This seems to indicate that the net income of the IPPs as a whole would decline in general with increasing target reserve %. A range of 10 to 12 % seems to be a reasonable level where the net income is relatively good and the reserve margin may cover the average effect

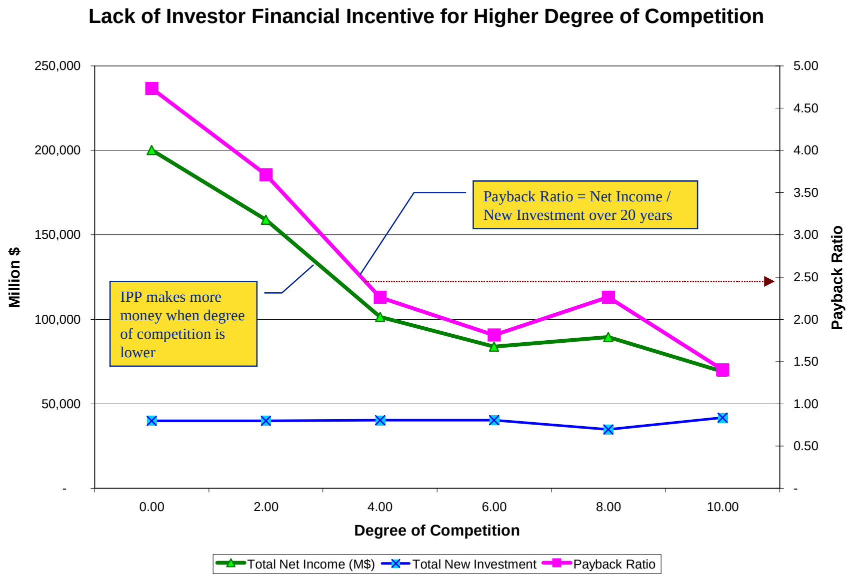


Figure 13 – Net Income and Payback Ratio for IPPs vs. Degree of Competition

of plant forced outages. Figure 12 seems to show that under the study's assumptions, there is no clear financial incentive for the IPPs to build surplus capacity.

To obtain some insight on the implication of the Competitive Degree, its effect on net income and the payback ratio for new investment is shown in Figure 13. When it is varied from zero to 10, the net income to the IPPs over 20 years drops by a factor of more than two. The payback ratio, which is the cumulative net income divided by the new investments, drops from 4.5 to 1.5. It should be noted that even at the base assumption of 10, the payback ratio is still reasonably profitable, at 1.5.

The LTPMS model also shows the cash flow of the IPPs as a whole over the 20-year simulation period. Figure 14 is such a graph for the base case. It indicates that the net income gain for the IPPs are highest during the period of capacity shortage and that negative net income may occur in those years when capacity is in surplus.

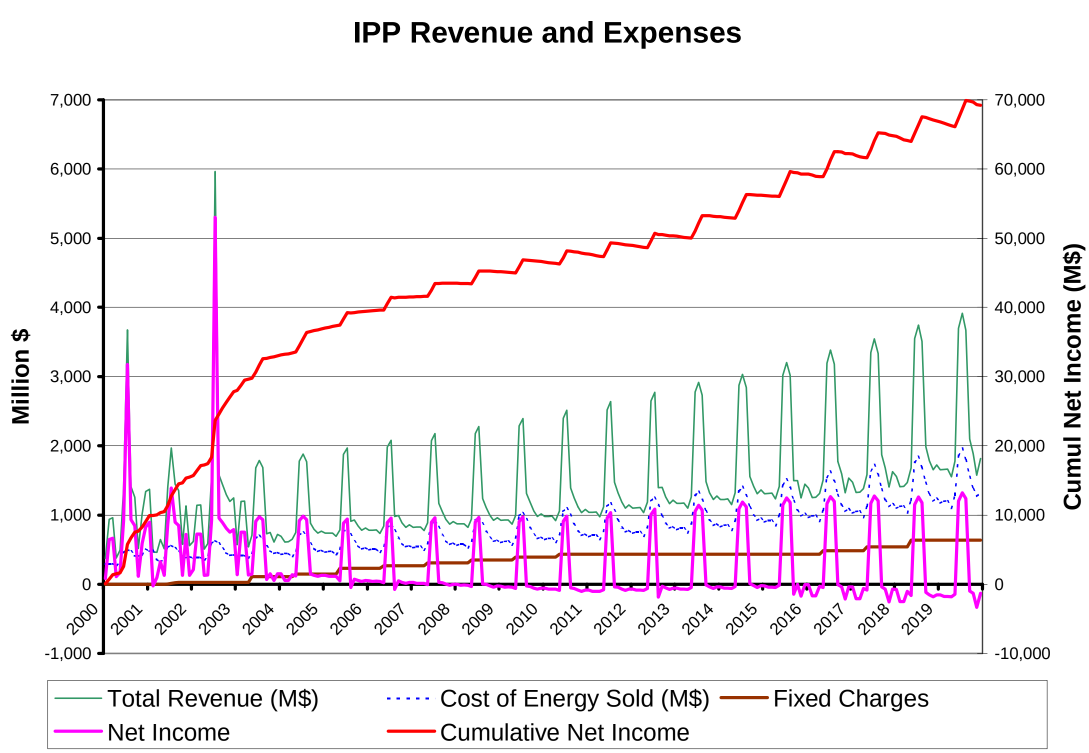


Figure 14 – IPP Revenue and Expenses Over 20 Years

Furthermore, the effect of the competitive degree on the customers can be seen in Figure 15. The costs of the electricity to the customers are added to the environmental cost of emissions and the social outage cost due to blackouts. On the right side axis of the graph is plotted the average reserve margin in the summer months over 20 years. It shows that with higher degree of competition, the average reserve margin increases from 18% to 30%. Note that this is the average over the possible boom and bust periods in 20 years. The total cost to customers and society decreases with higher degree of competition, and it is due to both the reduction of the electricity bill and the reduction in social outage costs.

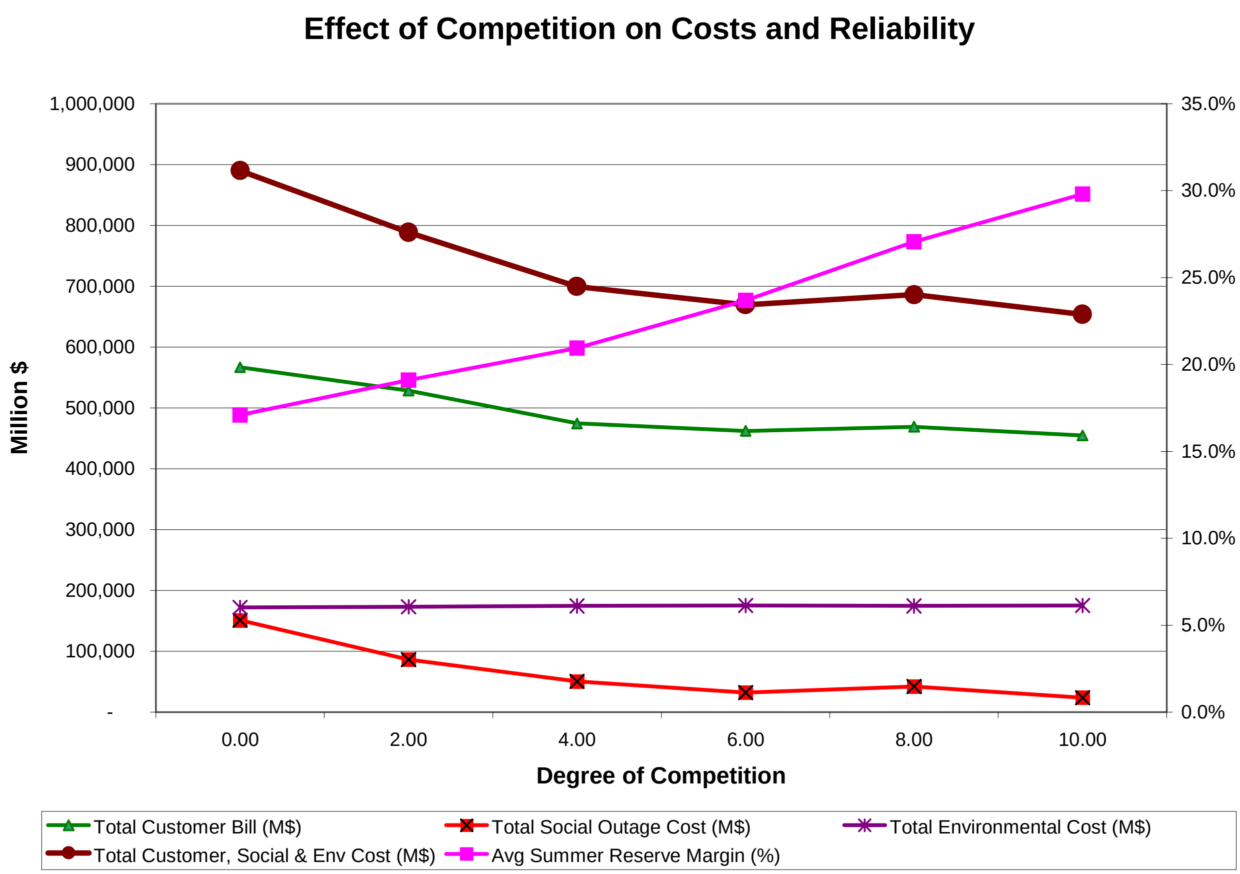


Figure 15 – Effect of Competition on Costs to Customers and Society

3. Effect of Demand Elasticity on Long-Term Price Volatility

From the base case simulations, which assumed no bilateral contracts, the total electricity rate to the customers exhibits high volatility from month to month. This is shown in Figure 16.

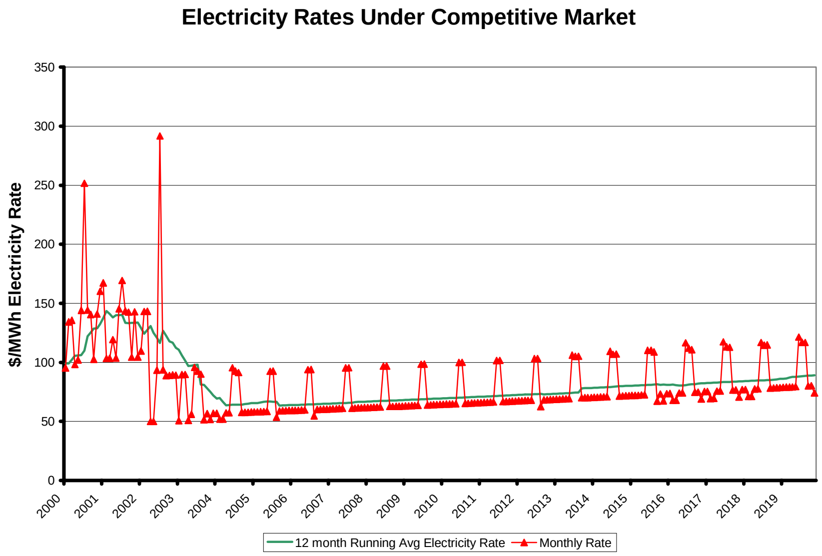


Figure 16 – Monthly and 12-month Average Electricity Rates Over 20 Years
No bilateral contracts, demand elasticity = 0.2, competitive degree = 10

The price spike in the summer of 2000 can be seen in Figure 16, which shows that the electricity rate (assuming no rate freeze) would have gone up from $100/MWh to about $250/MWh. The effect in the winter of 2000 was due to the natural gas price spike. Then the effect of demand elasticity and conservation showed itself in the summer of 2001 with a lower price spike. However, the simulation shows that in the summer of 2002, when the effect of demand elasticity relaxes and the load comes back up, a higher price spike may return. After that, sufficient new capacity would be on line to lower the market prices to a normal level. The graph of the 12-month running average electricity rate shows that if some form of rate averaging is applied, the customers would see less volatility on their electric bill. However, it should be noted that rate averaging diminishes the effect of demand elasticity and is counter to the concept of real-time pricing.

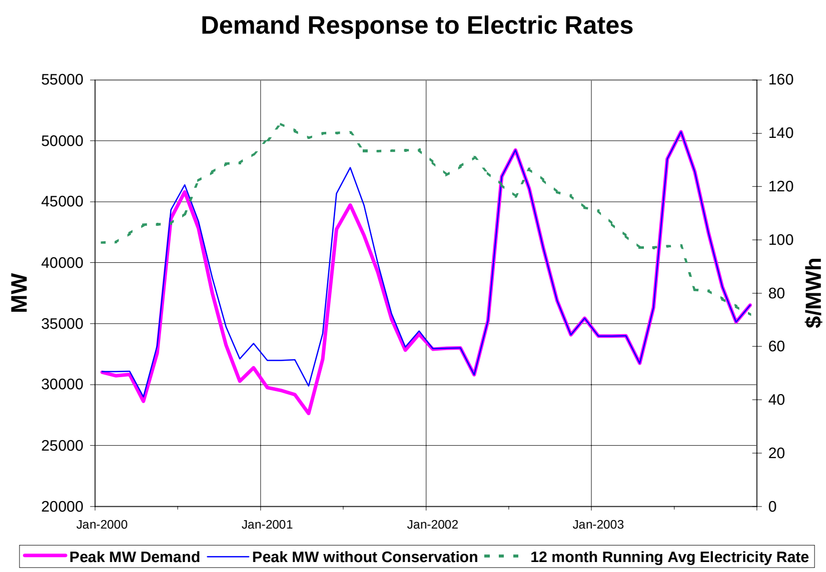


Figure 17 – Demand Response to Electric Rates at Demand Elasticity = 0.2

Figure 17 shows the monthly peak loads of the system from 2000 to 2003 under the assumption of 0.2 demand elasticity. The effect is to significantly reduce the customer demand from the winter of 2000 through the summer of 2001.

As a measure of price volatility, the LTPMS model computes the standard deviation of the monthly electric rate over the 20 year period. Two times the standard deviation plus the average rate is a reasonable number to use for measuring a upper range for the volatile price. To compare the effectiveness of bilateral contracts and demand elasticity on reducing price volatility, the LTPMS model was exercised in a set of sensitivity runs.

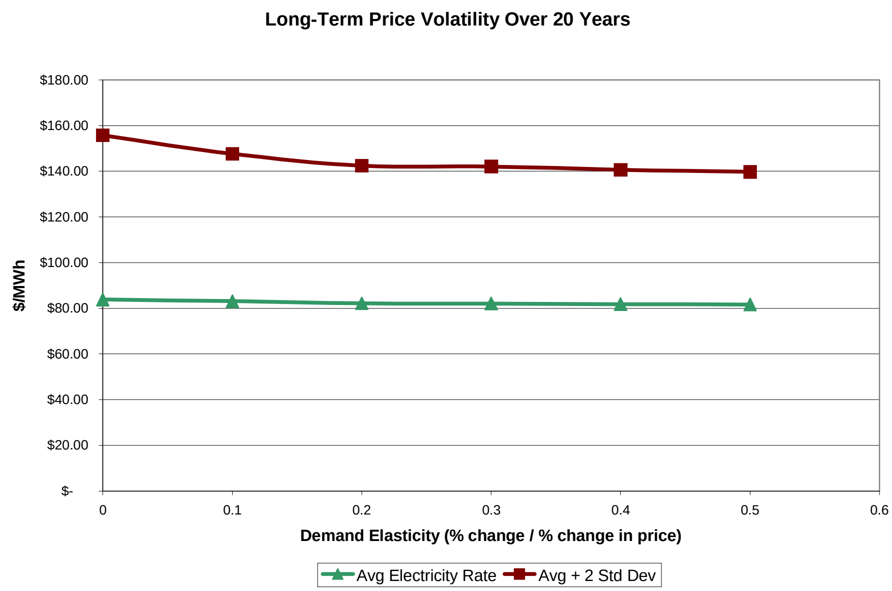
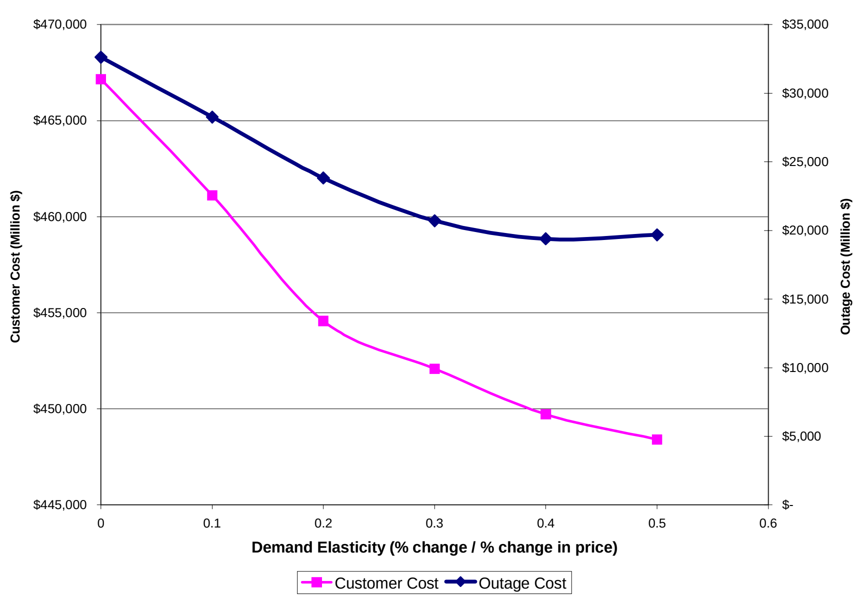


Figure 18 – Effect of Demand Elasticity on Long-Term Price Volatility

Figure 18 shows the effect of changing the assumption about demand elasticity on the average electricity rate and the upper range of the price over 20 years. Demand elasticity was changed from 0.0 to 0.5. It appears that demand elasticity itself has a significant effect on lowering long-term price volatility. Figure 18 seems to indicate that the upper range of the electric rate over the long term can be reduced from about $160/MWh to $140/MWh when demand elasticity is allowed to take effect by influencing the demand consumption due to the year-to-year rate increase. It may also be attributed to the effect of an emergency appeal for conservation. This result does not assume the application of real-time pricing which should further reduce the short term price volatility.



In terms of the effect of demand elasticity on the total cost to the customers and to society over 20 years, Figure 19 shows that the initial effectiveness of demand elasticity on both the customer cost and the social costs is very high.

Figure 19 – Effect of Demand Elasticity on Customer and Total Costs

4. Effect of Bilateral Contracts on Long-Term Price Volatility

A set of sensitivity runs was made by changing the assumption on the percentage of energy under bilateral contracts, ranging from 10% to 100%. Recalling that the bilateral contracts were assumed to start at an energy rate of $65/MWh for 10 years, followed by $87/MWh for 10 years, increasing the amount of energy under bilateral contracts also has the effect of changing the average electricity rate to the customers.

Figure 20 shows the effect on the 20-year average electricity rate and the price volatility index represented by twice the standard deviation of monthly rates over the 20 year period. The average electricity rate goes up with increasing percentage of bilateral contracts, while the price volatility index goes down with increasing bilateral contracts. This supports the intuitive belief that bilateral contracts reduce long term price volatility. The increase in average rate is an indication that the price of the bilateral contract is higher than market price.

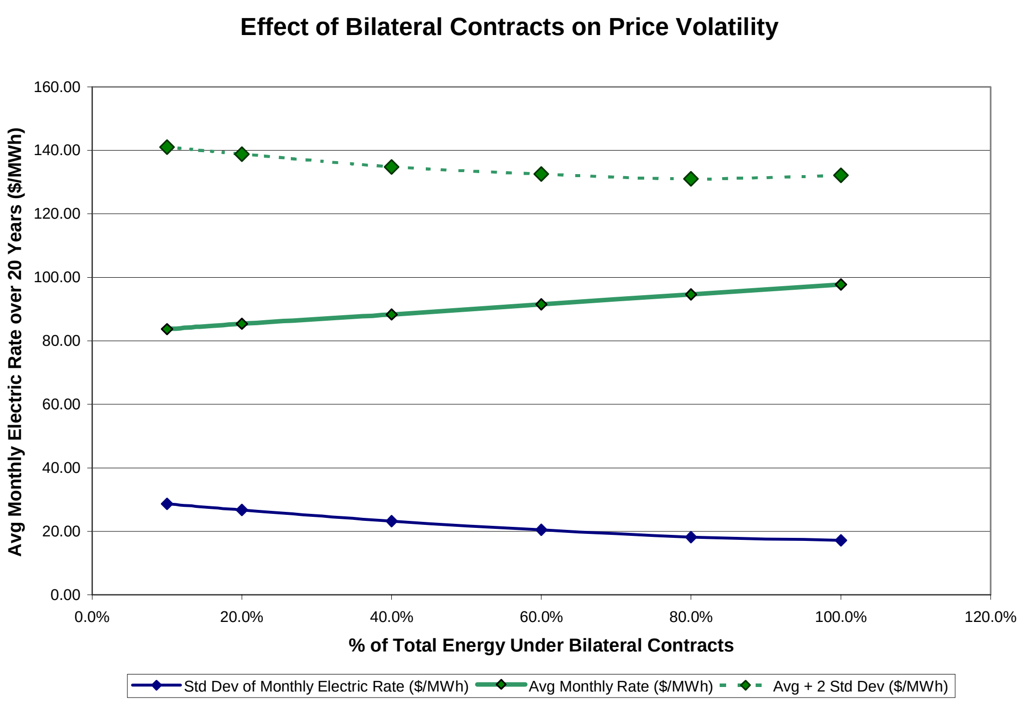


Figure 20 – Effect of Bilateral Contracts on Price Volatility

5. Exploring a Hybrid Power Market

A set of runs were made using the hybrid market version of the LTPMS model. The base case assumption is that the power authority (PA) will build or procure combustion turbine capacities equal to 30% of the annual MW load growth, if the current summer's B/C ratio for gas-fired generation capacity, as seen by the IPPs drop below 1.0. With this assumption, the 20-year simulation shows a demand/resource picture shown in Figure 21.

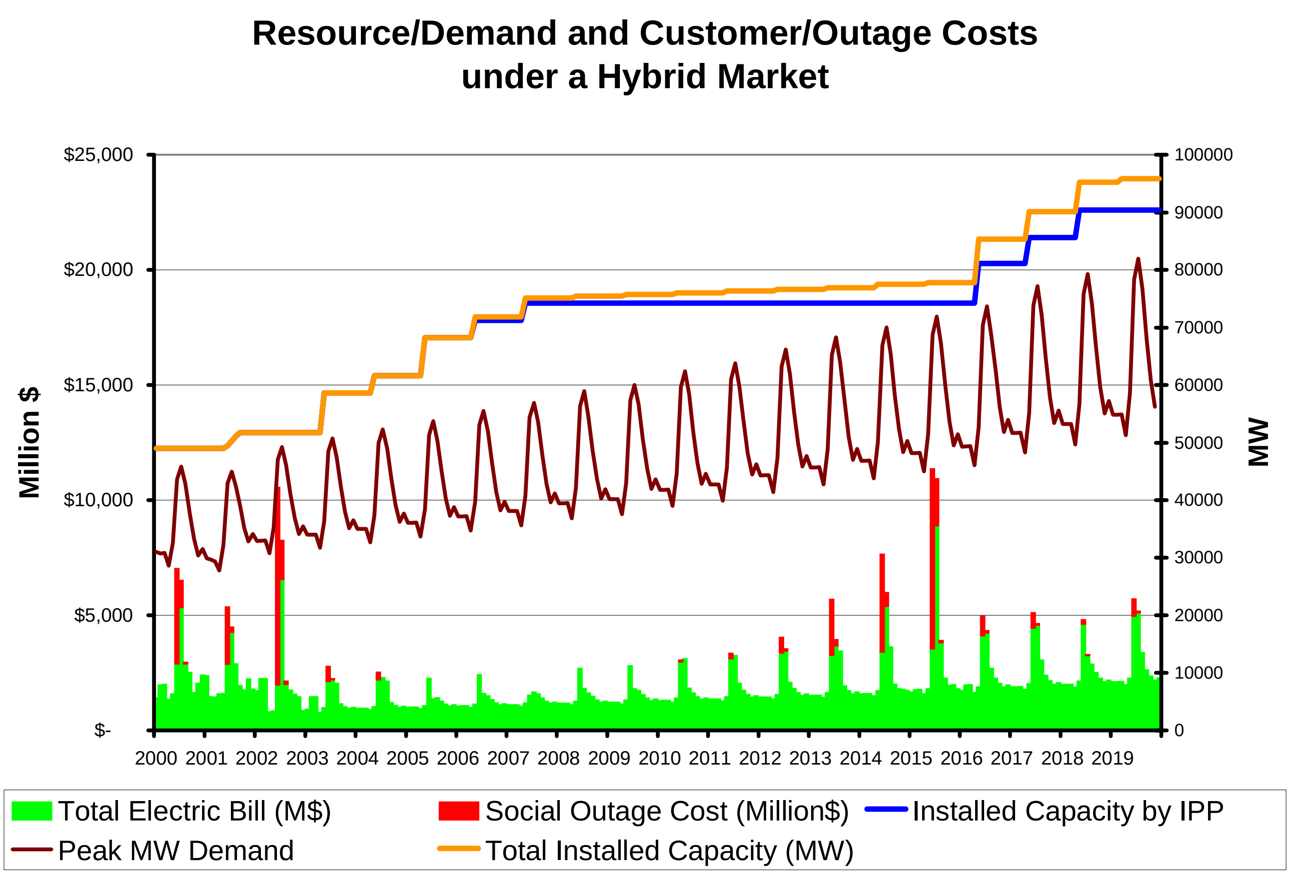


Figure 21 – Simulation Results of a Hybrid Market with PA Supplying 30% of the Annual Load Growth on a Contrarian Principle

It shows that the PA will only start putting new CT on line in 2006. From 2008 to 2015, no new capacity is added by the IPPs. Only when the reserve margin falls to the point when the B/C ratio starts coming above 1.0 during the summer months would the IPP resume building new capacity. Comparing Figure 21 to Figure 10 shows that without the participation of the PA on the supply side, the IPPs would continue to have new capacity coming on line through 2011 and resume having new capacities on line in 2016. In other words, the participation of the PA on the supply side may hasten the decisions of the IPPs not to build and cause the hiatus of their new capacity increasing from 5 years to 8 years.

The effect of changing the decision rule of the PA on the overall supply picture is studied with the LTPMS model. It is summarized in Figure 22a and 22b.

As shown in Figure 22a and 22b, the total capacity built by the PA increases with its decision rule to build greater percentages of the annual load growth on the contrarian principle. However, the total capacity built by the IPPs decreases rapidly when that percentage goes above 30%. In fact, because the total capacity built by the IPPs include the amount built before the PA starts doing it in 2006, on more detailed examination, the IPPs stop building altogether when the PA's decision rule is 60% or above.

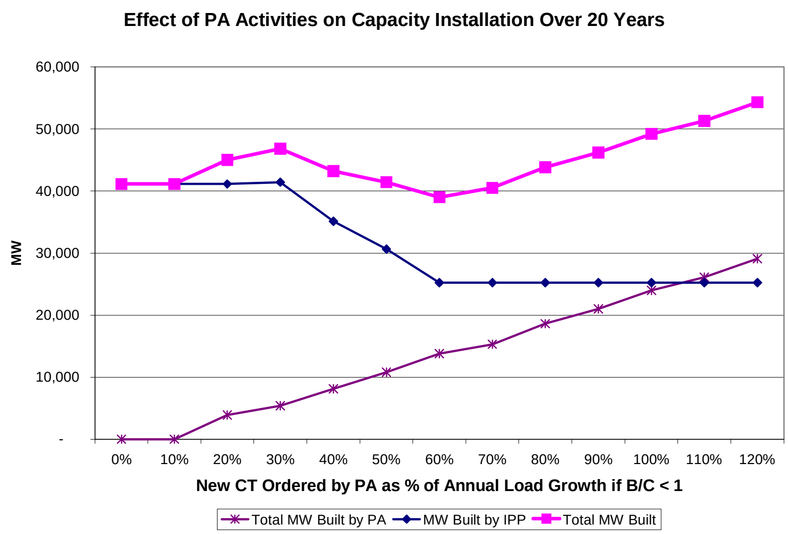


Figure 22a – Effect of PA's Supply Side Decision Rule on Capacity Additions Over 20 Years

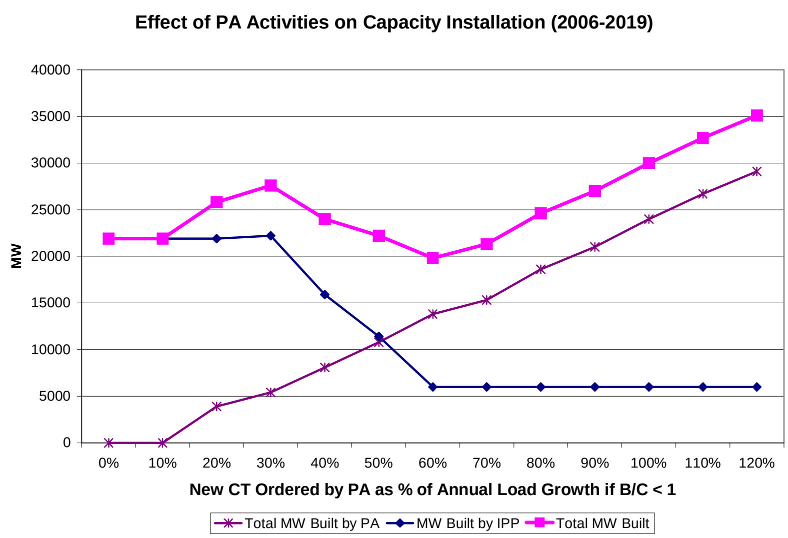


Figure 22b – Effect of PA's Supply Side Decision Rule on Capacity Additions (2006-2019)

The cash flow of the IPPs as a whole, under the base case assumption of the PA building for 100% of annual load growth, is shown in Figure 23.

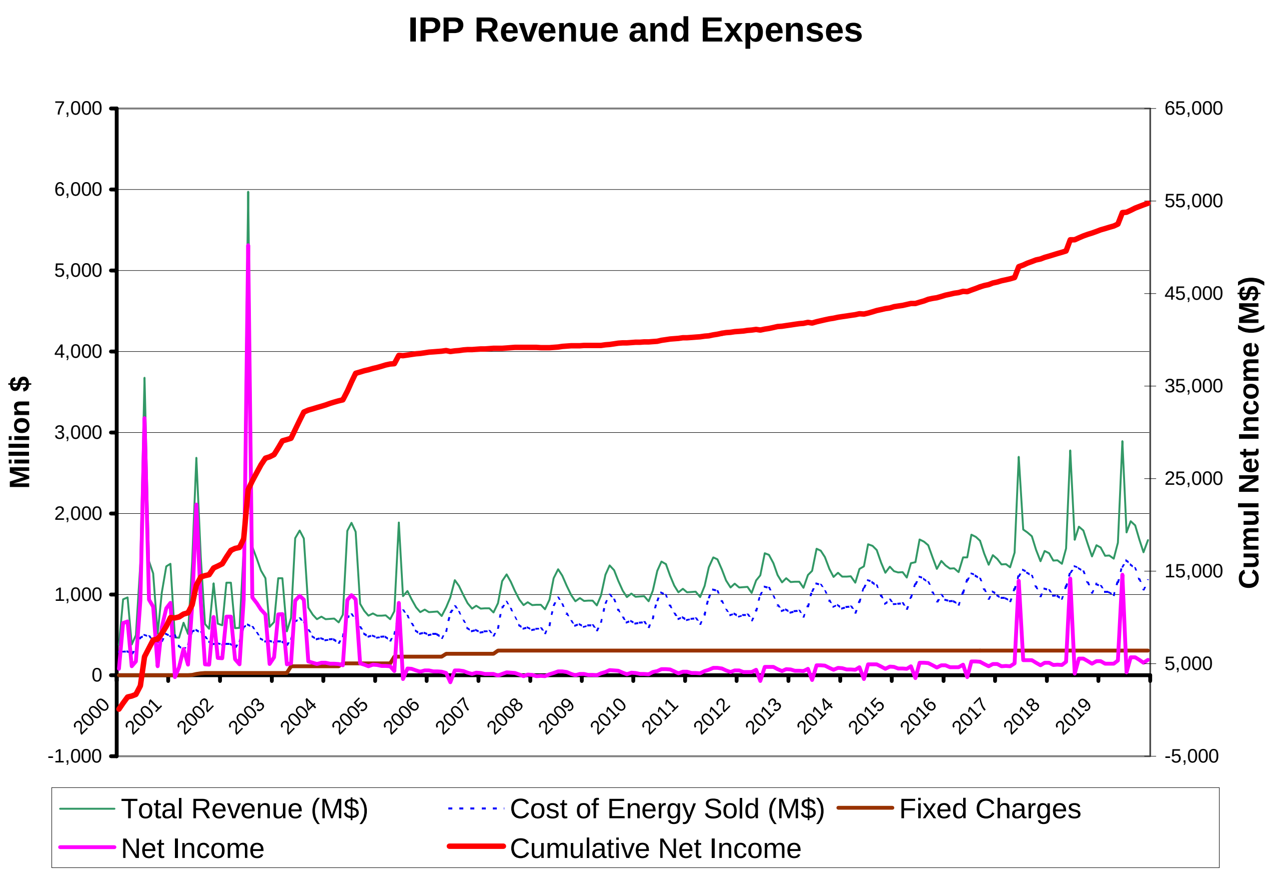


Figure 23 – IPP Revenue, Expenses and Net Income under Hybrid Market, with PA building for 100% of Annual Load Growth

Figure 23 shows that the net income of the IPPs may go into a breakeven situation when the PA enters the supply side at a high level of building capacity equal to 100% of each year's load growth. Because that does not provide for reserve margin, when the initial surplus reserve is depleted, market prices start increasing and the IPPs will receive moderate amounts of net income. In this chart, note that the net income is the total revenue minus cost of energy sold minus fixed charges, which includes an allowance for 15% return on investment.

Participation by the PA on the supply side does not appear to have much effect on long term price volatility. As shown in Figure 24, the upper range of the average electricity rate did not decrease over the entire range of the PA decision rule. In fact, there appears to be a range between 40% and 70% where the price volatility may increase. This was due to the phenomenon when the IPPs would withdraw from building new capacity, and the PA was not building enough for the whole system's load growth, causing capacity shortages in some years.

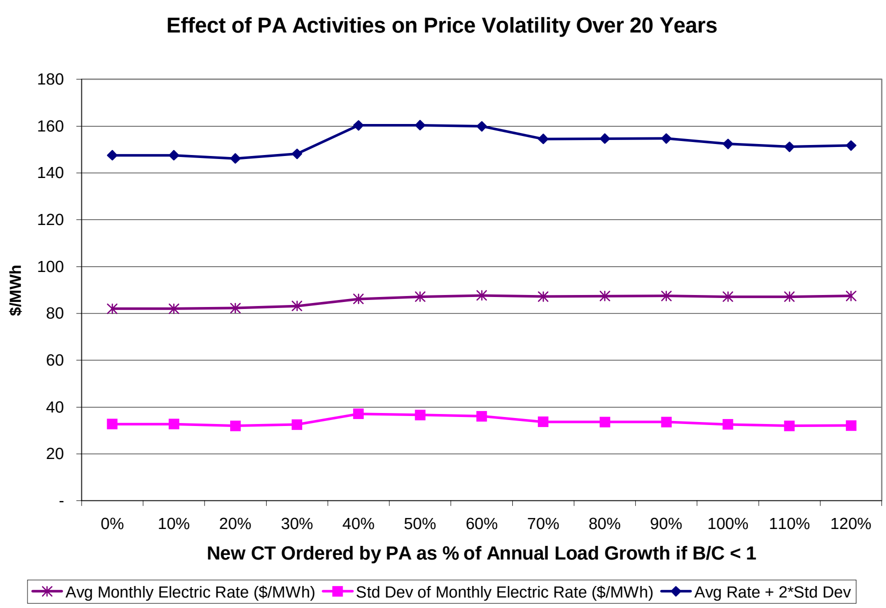


Figure 24 – Electricity Rates Under a Hybrid Market
No bilateral contracts, demand elasticity = 0.2, competitive degree = 10

Figure 25 shows the effect of the PA activities on the costs from 2006 to 2019, which are the years when the PA activities started to have some effect. It shows that when the PA builds more than 80% of the load growth each year, the social outage costs become insignificant. However, between 40% and 70%, the total costs (customer bill plus environmental and social outage costs) become higher. Part of the reason is because the decision rule that was modeled allows the PA to build only combustion turbines. Thus, it is not realistic in the high % range. In reality, if the PA was to build more than 40% of the load growth, it would likely include non-peaking generation in its mix of portfolio, which may reduce the customer bill, as compared to building only combustion turbines for an extended period of time.

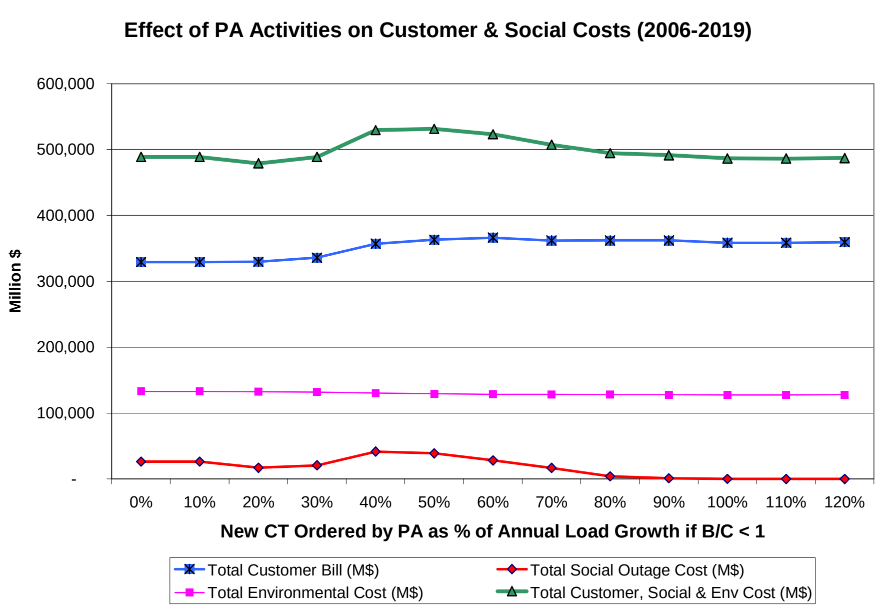


Figure 25 – Effect of PA Activities on Customer and Social Costs (2006-2019)

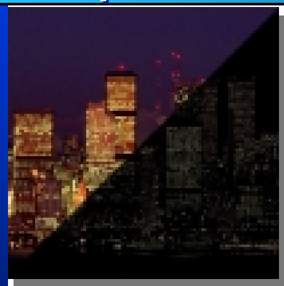
Figure 24 and Figure 25 together seem to indicate a risk to the health of the power market if the PA's participation in the supply side is not carefully studied. At a low level of participation, the benefit seems to be mildly beneficial. In the middle range, it raises the risk that the IPPs will reduce their building of new capacity. Beyond that mid-range, it would seem that the IPPs may withdraw from building new capacity.

In summary, the LTPMS model has demonstrated its ability to simulate a competitive power market and a hybrid power market over a 20-year period. Its results, although still preliminary, provide insights into how different power market structures and demand elasticity may affect long term price volatility. Furthermore, it shows interesting insights as to how a Power Authority's decision rule to build peaking capacity in a contrarian principle, i.e., when the market decides not to build, may interact with the market for competitive generation capacity. The purpose of this paper is to present objective results from this model and not to propose or advocate public policies.

References

1. Stephen Lee, "Lessons Learned from the California Power Crisis," a companion discussion paper, November 1, 2001.
2. Andy Ford, "Simulating Patterns of Power Plant Construction with the CEC Model - Summary Report to the California Energy Commission," November 14, 2000.
3. Andy Ford, "Waiting for the Boom: A Simulation Study of Power Plant Construction in California" (to appear in Energy Policy, Vol. 29, p. 847-869), March 2001.
4. Robert J. Thomas, et al, "A Comparison of the Results of Three Auction Experiments," Cornell University, December 2000.

Long-Term Power Market Simulations and Comparisons Of Alternative Structures



Stephen Lee
Area Manager
Grid Operations and Planning
EPRI

CEC Workshop
November 7, 2001
Sacramento, CA

Disclaimer

- Results herein are for discussion purposes only and do not imply any policy recommendation by the author or EPRI
- In the technical field of “Modeling,” it is necessary to make simplifying assumptions. Therefore results from models must be interpreted carefully, subject to a full understanding of all the assumptions, simplifications and limitations involved.

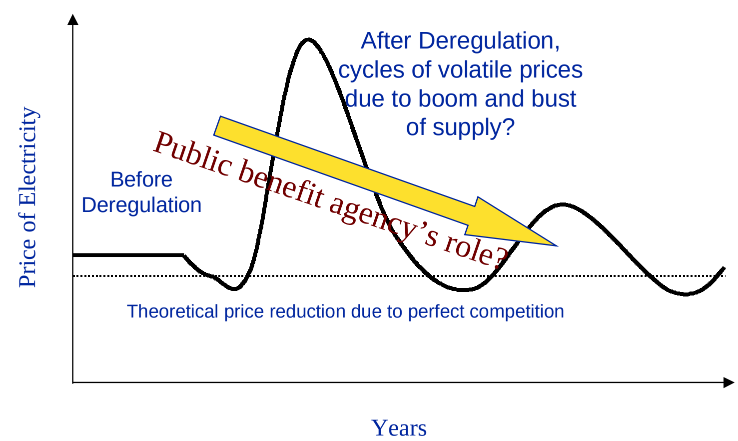
Outline of Presentation

- Objectives of Long-term Power Market Simulation
- Assumptions and Limitations of the Simulation Model
- General Results of Simulations
- Some Movies of Simulations
- Some Results on Long-term Price Volatility
  - effect of demand elasticity
  - effect of bilateral contracts
- Results on 1 hypothetical supply rule by the Power Authority

Objectives of Long-term Power Market Simulation

- Understand how a competitive power market works in the long term
- Study the potential long-term price volatility and boom-bust cycles of the power market
- Study alternative market designs and the role of a Power Authority to complement the market
- Compare the long-term impacts of various market structures on the end-users and society

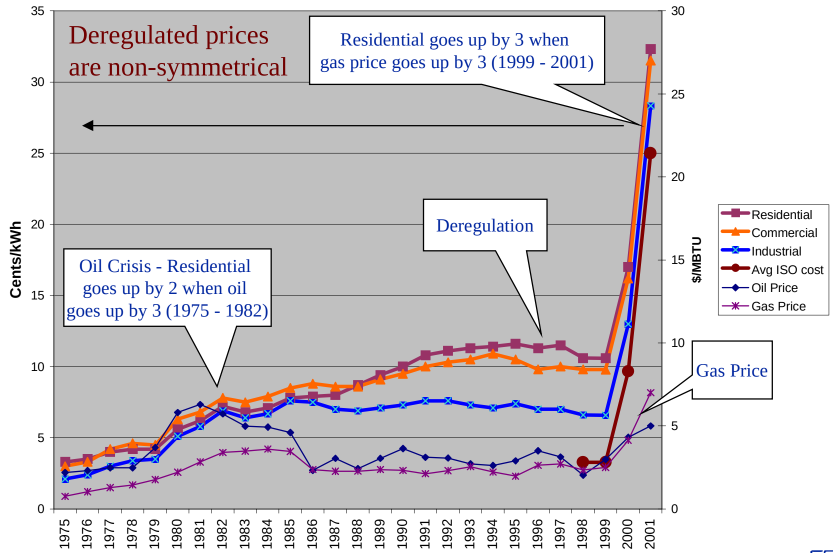
Road to Perfect Competition Or a Hybrid Power Market?



Characteristics of Competitive Market

- Profit incentive drives innovation and efficiency
- Attracts investment and new participants
- Liable to boom-bust cycles because of the focus on maximum profits and fast payback on investments
- Prices are non-symmetrical. Big spikes may occur during supply shortage
- Does not consider all public benefits needed by society

Volatile Prices in California After Deregulation



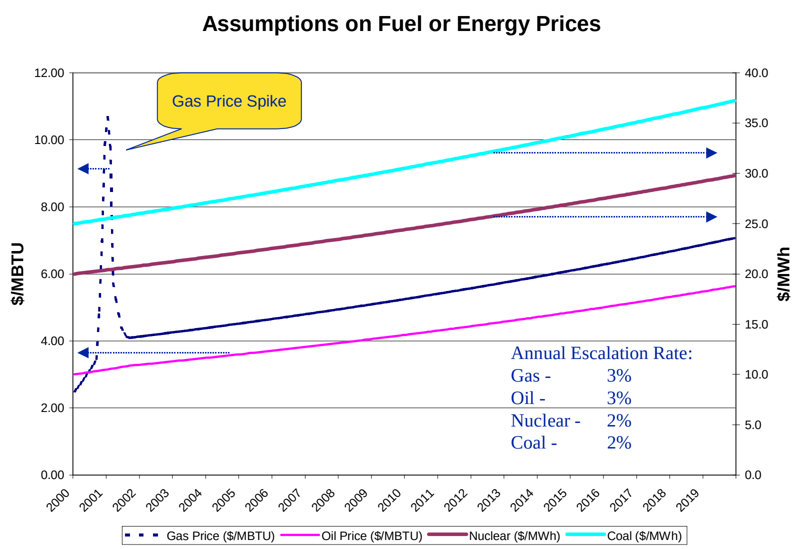
General Assumptions and Limitations of the Simulation Model

- Uses Microsoft Excel - great for testing ideas
- Simulation is monthly over 20 years (2000-2019)
- Entities which are modeled:
  - Grid Operator, Independent Power Producers (IPP), Customers
  - Power Authority in the hybrid scenario
- System size and generation mix approximate the state of California
- Load growth is 3% per year

Assumed Wholesale Market Structure

- Grid Operator (ISO)
- All generators submit bids to a central market
- The clearing price applies to all successful bidders, subject to a price cap
- Price cap is $250/MWh in 2000, escalating at 3% per year
- No capacity payments or capacity market
- A user-input % of the entire generation mix may be placed under bilateral contracts
  - at $65/MWh for 10 year (fixed price) and renewed for 10 years at $87/MWh (3 % annual inflation adjustment)

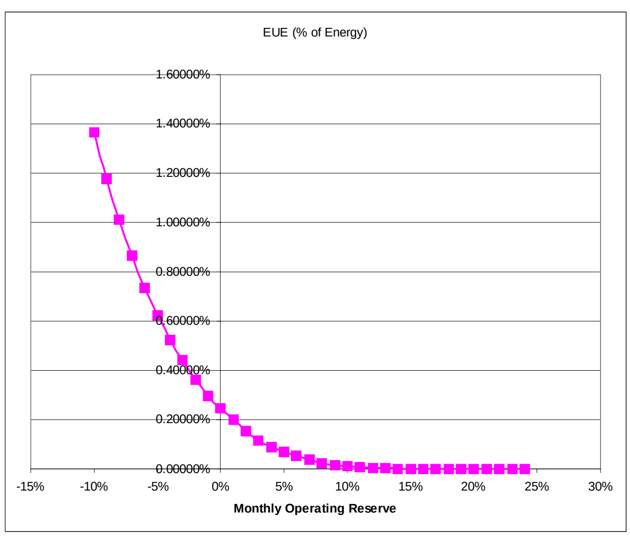
Fuel and Energy Price Assumptions



Power System Model

- Generators are scheduled for maintenance outages away from June, July, Aug, and Sept
- Generators have an average forced outage rate of 8%
- Hydro generation is based on average water condition
- Random forced outages will cause Expected Unserved Energy (EUE), using a curve computed from a probabilistic method

Expected Unserved Energy Model



Assumptions on Societal Impacts

- Emission costs:
  - NOX $30/lb (in CA, the range was $5-40/lb)
  - SO_2 $250/ton
  - CO_2 $1/ton
- Society's Economic Cost of Expected Unserved Energy (modeled as a function of reserve margin) = $100,000 MWh (this results in $7 billion for 2000-2001 in the model)

Assumptions about IPPs

- Market structure requires the IPPs to make enough operating income from its energy sales to recover the investment cost of new or existing power plants
- Assumed IPP decision rule to build new plants
  - If the B/C ratio of a new plant > 1.0, existing generating companies in the market will order new units to meet a projected target reserve margin of 12% (may be varied in the model) which will be online 3 years later
  - Power plants will also be built by new competitors, without considering the target reserve (base case = 10 x 300 MW units, no. can be varied to model different degrees of market competition)

Assumptions about IPPs

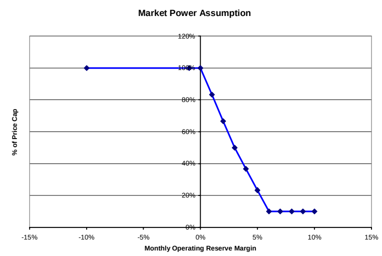
- Benefit (in \$/MWh) = market clearing price - variable operating cost of a gas-fired unit
- Cost (in \$/MWh) = levelized annual fixed charge of a gas-fired unit prorated over the MWh expected to be generated
- B/C ratio is computed for each month, and the summer B/C ratio drives the IPP decision
- New capacity is 80% from gas-fired combined cycle units and 20% gas turbines
- Avg capital cost of gas-fired generation = \$600/kW
- IPP's desired rate of return = 15% (+ 3% other fixed costs ➔ 18.23% fixed charge rate)

Assumptions about Power Authority

- PA's approved rate of return = 12% (+ 3% other fixed costs ➔ 15.41% fixed charge rate)
- In the Hybrid Market scenarios, the PA adopts a "Contrarian" principle, to complement the market decision to build or not to build
- If as seen by the IPPs, the B/C ratio is <1 (uneconomical), then the PA will order new combustion turbines to be on-line in 2 years, at an amount equal to X% of the annual peak load growth.
- X is varied by the model between 0% and 120%

Assumptions about Market Power

- Market power is a function of reserve margin



IPPs assumed to bid at 105% of marginal energy cost at the minimum. This curve is a macro relationship and does not imply that IPPs use the reserve margin to set their bid price.

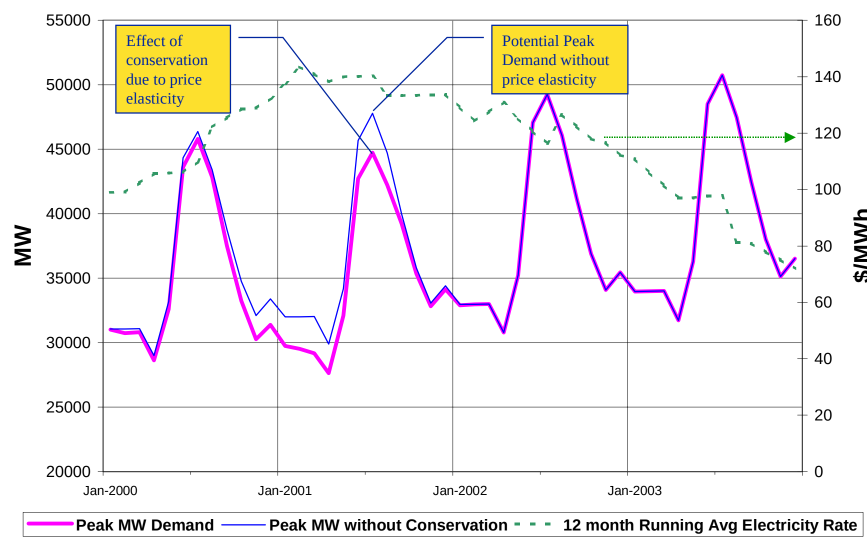
- E.g., when reserve is at 3%, bid price = at least 50% of price cap
- When system reserve is negative, bid price = 100% of price cap (in reality, supply shortage may occur)
- These two factors combined benchmarked well to year 2000-2001 experience in California

Assumptions about Customers

- A 12-month running average total electricity rate is assumed to be paid by the customers
- Customers respond to the running average electricity rates with demand elasticity equal to 0.2% for each 1% increase in rates from 12 months prior
- The demand elasticity was varied from 0% to 0.5% for sensitivity analysis
- No real-time pricing is simulated
- The effect of 0.2% demand elasticity in 2001 was to reduce peak demand by about 3000 MW

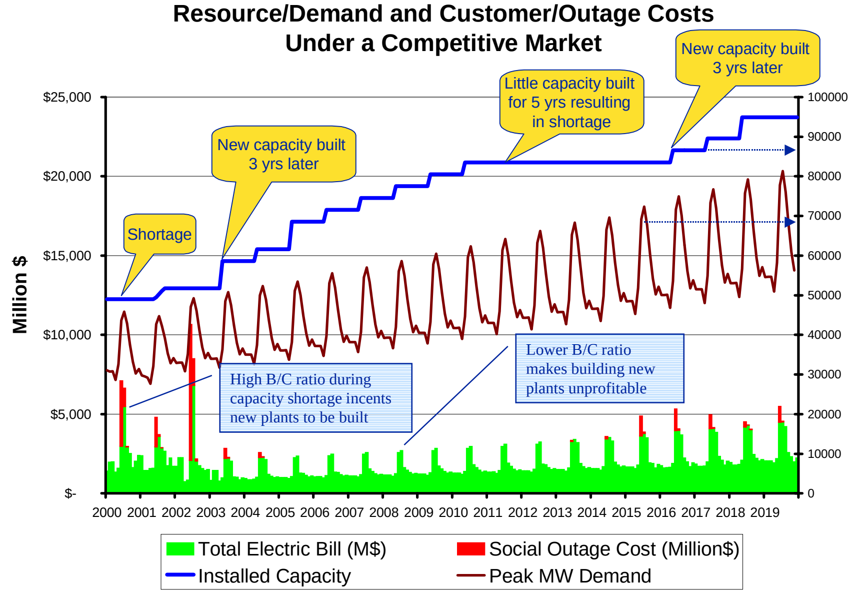
Demand Response

Demand Response to Electric Rates

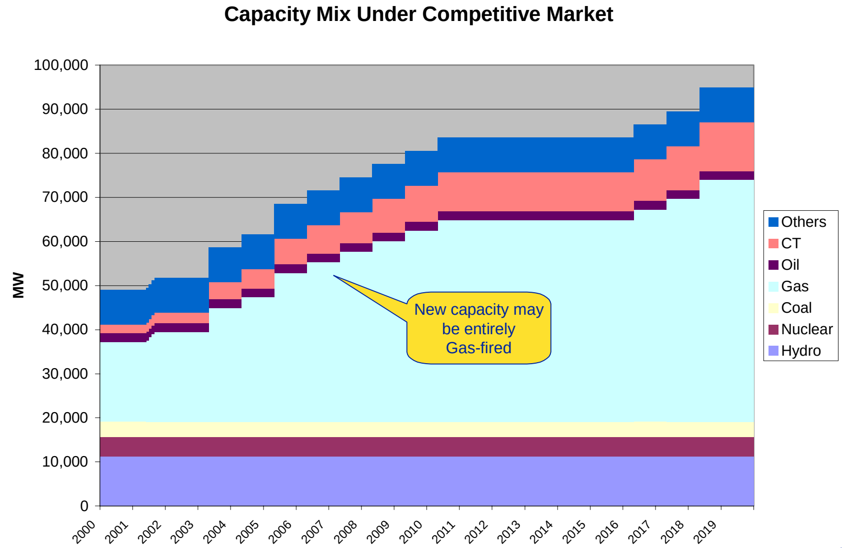


Simulation of California Market

Resource/Demand and Customer/Outage Costs Under a Competitive Market

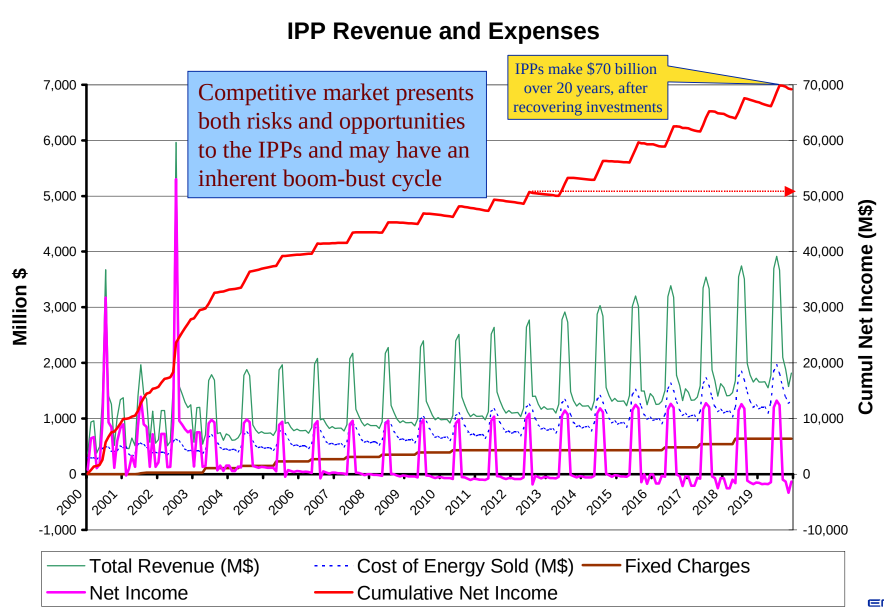


Generation Capacity Mix Under Competitive Market



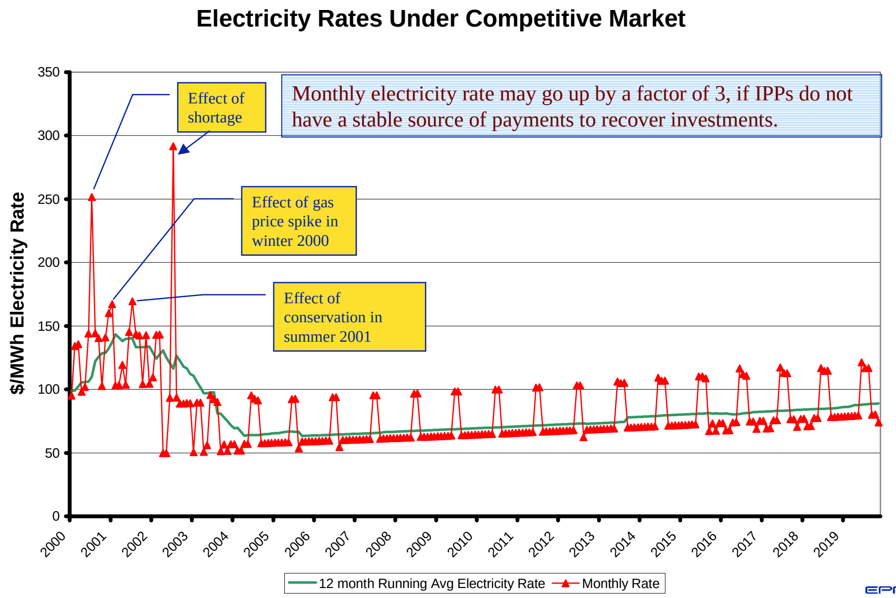
21

Cash Flow of IPPs



22

Electric Rates



Movie of IPP Target Reserve %

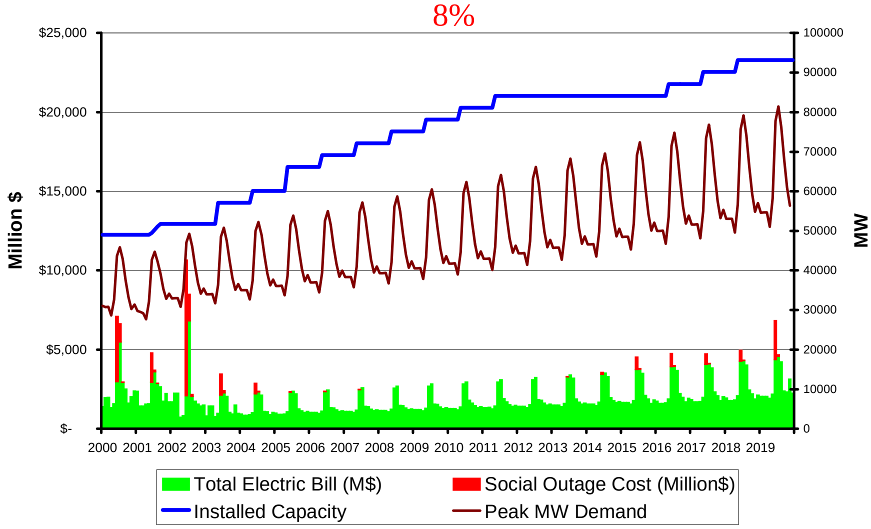
Demand elasticity, Target reserve %, Competitive degree

0.2,8%,10

0.2, 8%, 10

Resource/Demand and Customer/Outage Costs Under a Competitive Market

8%



25

EPRI

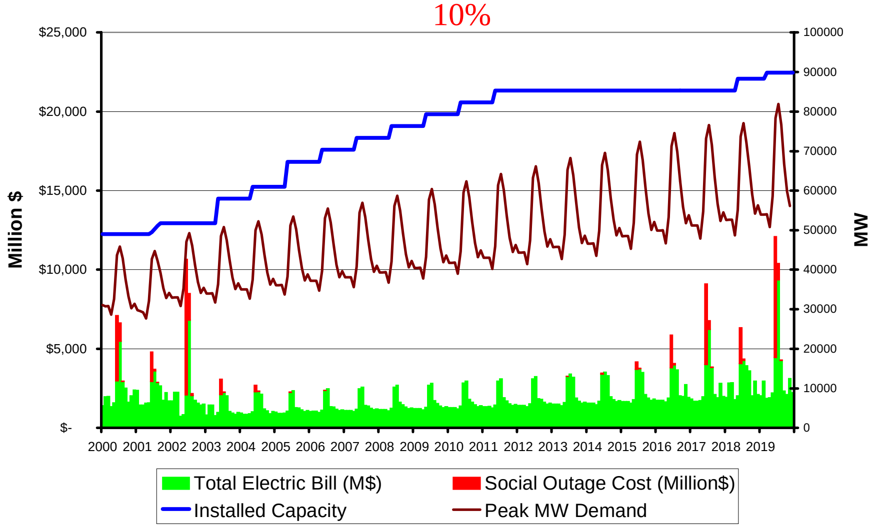
Demand elasticity, Target reserve %, Competitive degree

0.2,10%,10

0.2, 10%, 10

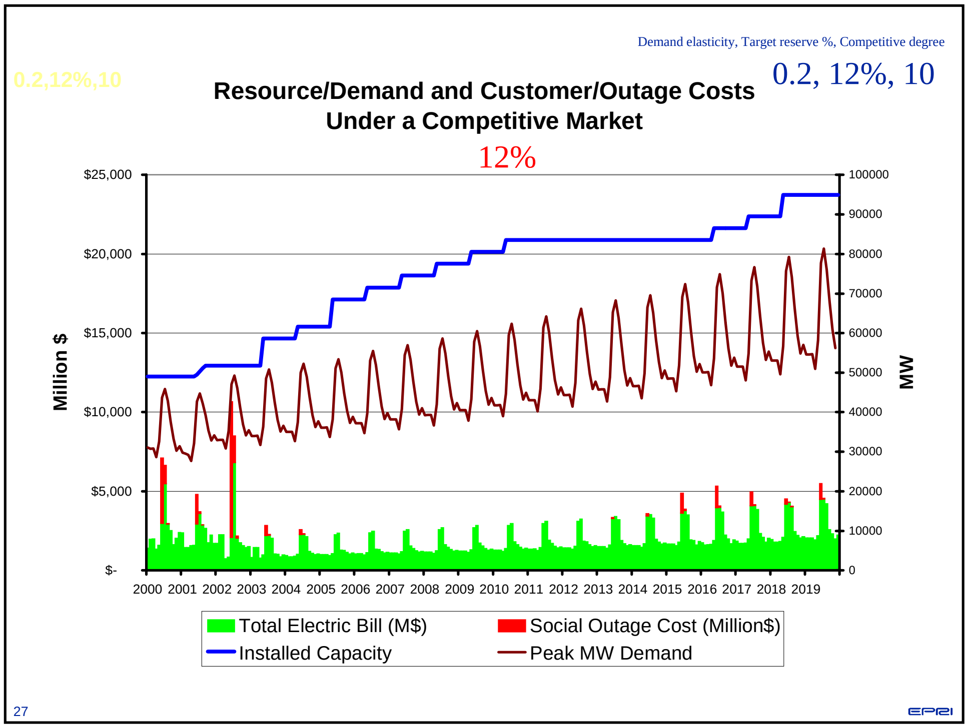
Resource/Demand and Customer/Outage Costs Under a Competitive Market

10%

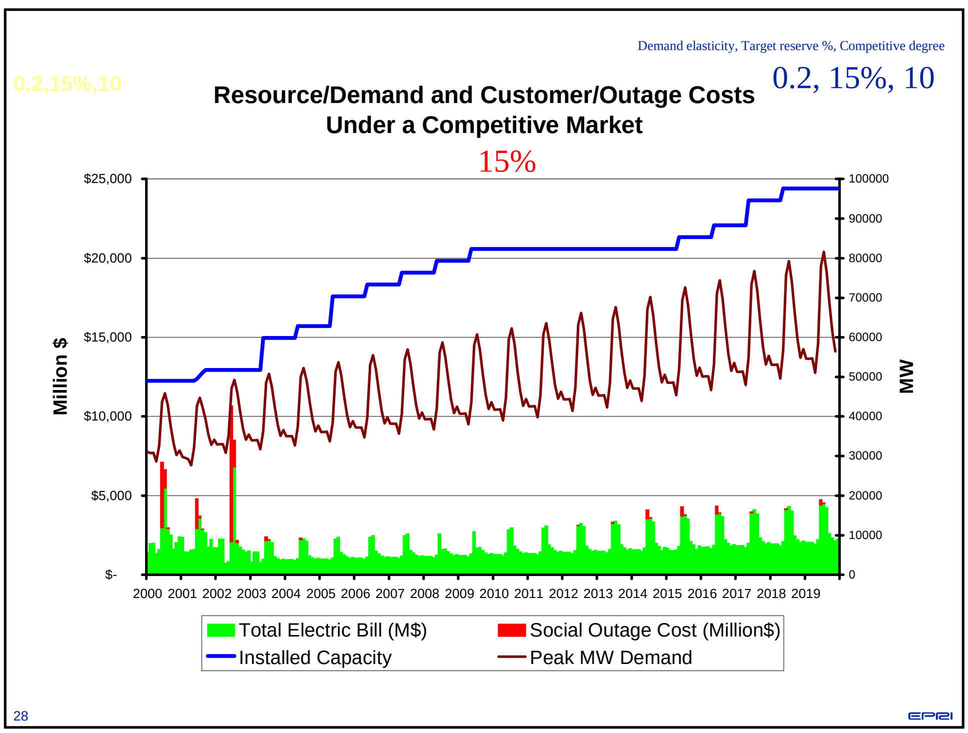


26

EPRI

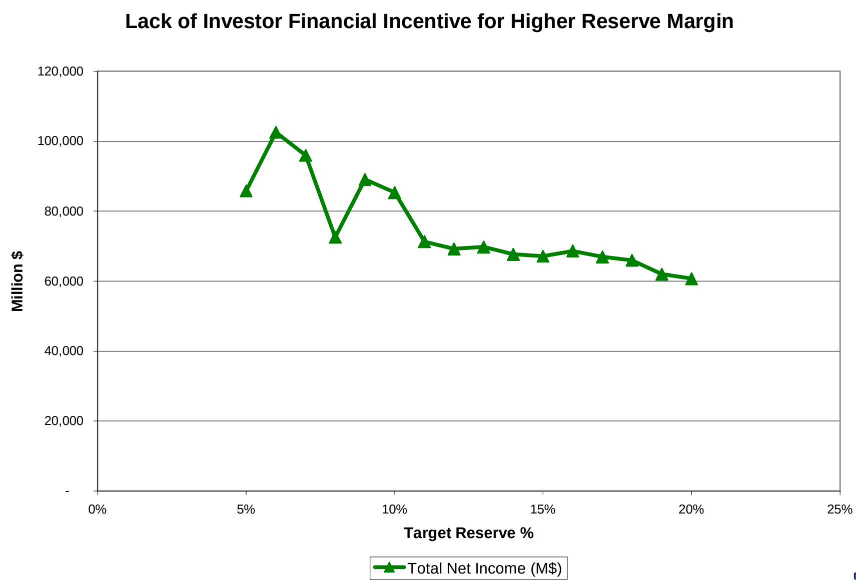


Demand elasticity, Target reserve %, Competitive degree
0.2,12%,10
0.2, 12%, 10
Resource/Demand and Customer/Outage Costs
Under a Competitive Market
12%
Million $
MW
$25,000
$20,000
$15,000
$10,000
$5,000
$-
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity
Peak MW Demand
27
EPRI



Demand elasticity, Target reserve %, Competitive degree
0.2,15%,10
0.2, 15%, 10
Resource/Demand and Customer/Outage Costs
Under a Competitive Market
15%
Million $
MW
$25,000
$20,000
$15,000
$10,000
$5,000
$-
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity
Peak MW Demand
28
EPRI

Effect of Target Reserve on IPP Net Income



Movie of Competitive Degrees

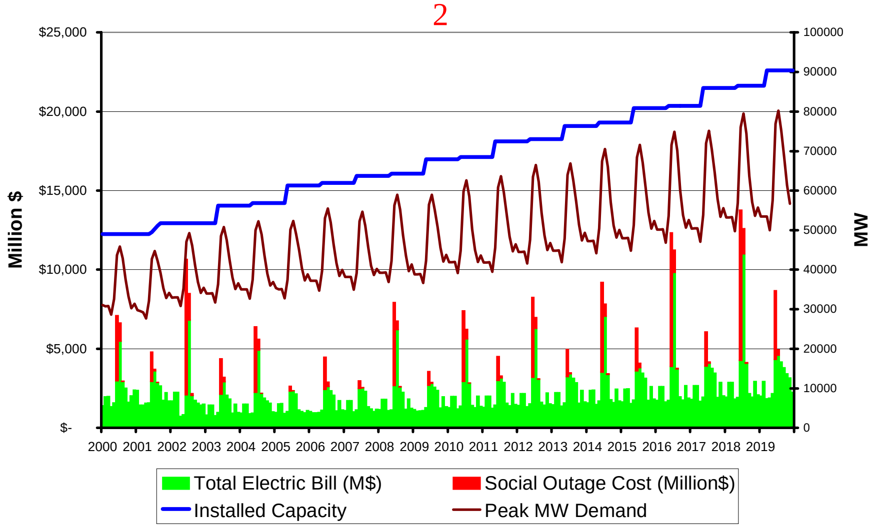
Demand elasticity, Target reserve %, Competitive degree

0.2,12%,2

0.2, 12%, 2

Resource/Demand and Customer/Outage Costs Under a Competitive Market

2



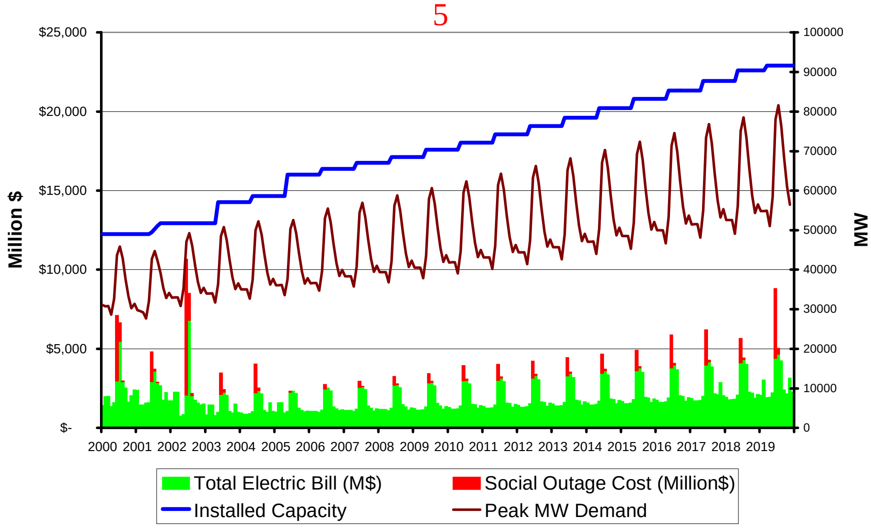
Demand elasticity, Target reserve %, Competitive degree

0.2,12%,5

0.2, 12%, 5

Resource/Demand and Customer/Outage Costs Under a Competitive Market

5



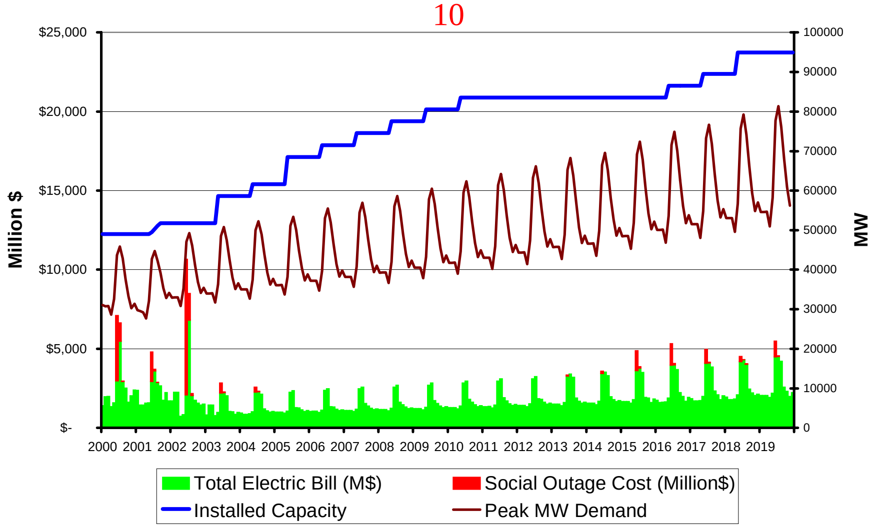
Demand elasticity, Target reserve %, Competitive degree

0.2,12%,10

0.2, 12%, 10

Resource/Demand and Customer/Outage Costs Under a Competitive Market

10



33

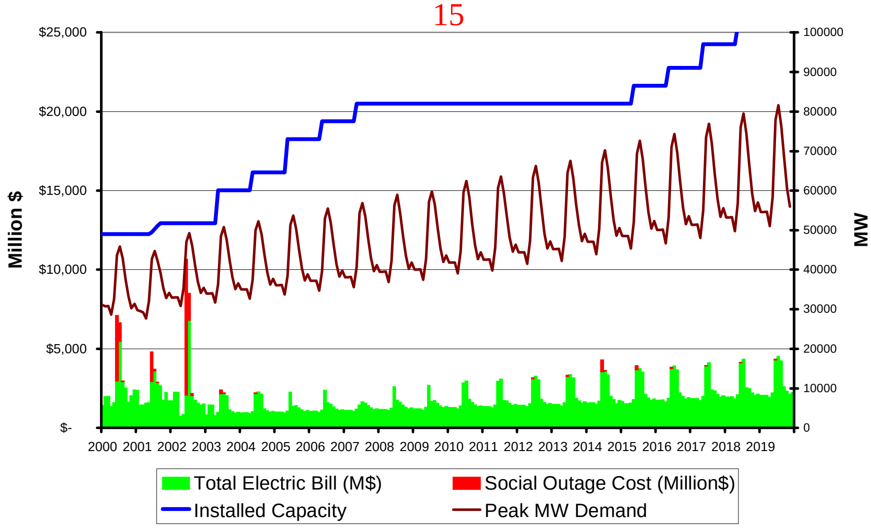
Demand elasticity, Target reserve %, Competitive degree

0.2,12%,15

0.2, 12%, 15

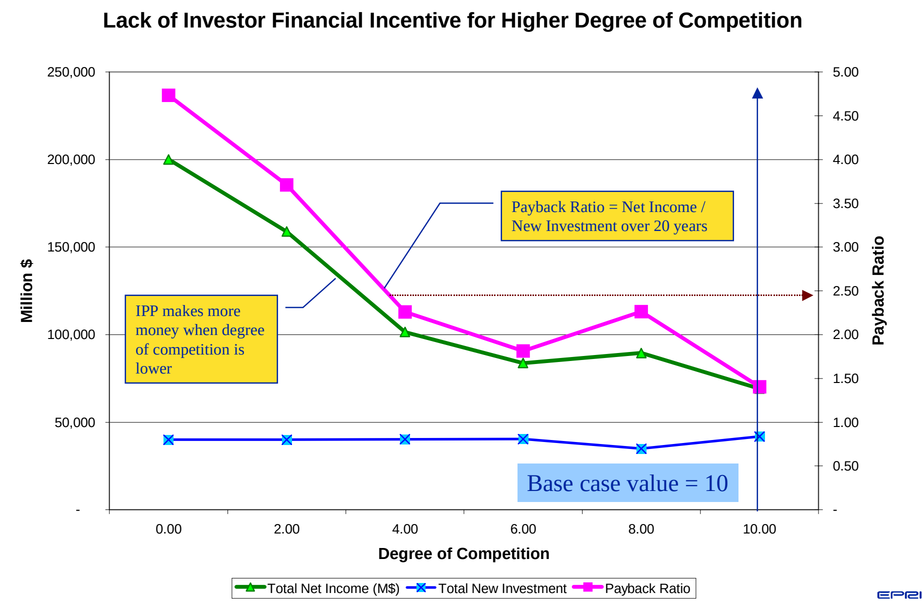
Resource/Demand and Customer/Outage Costs Under a Competitive Market

15



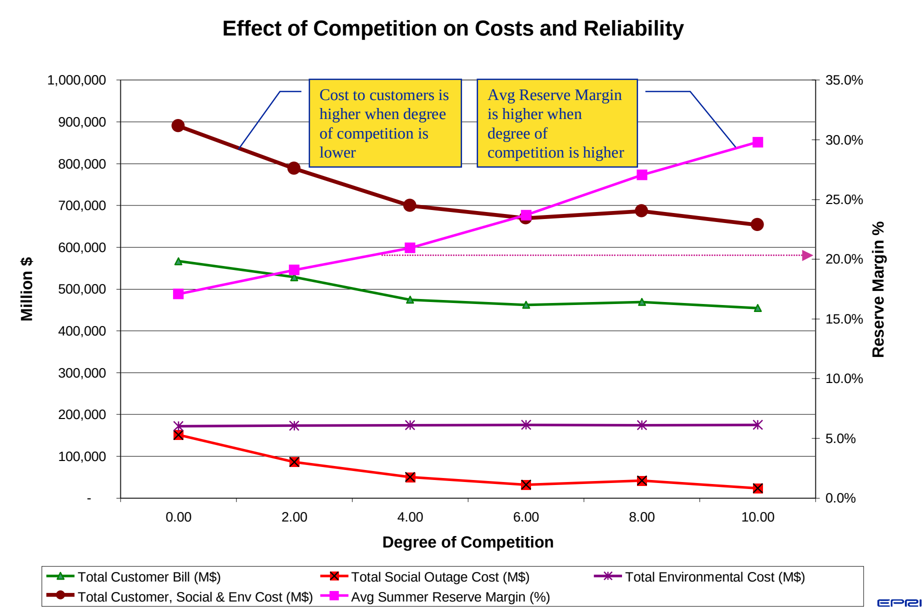
34

Effect of Competitive Degree



35

Effect of Competitive Degree



36

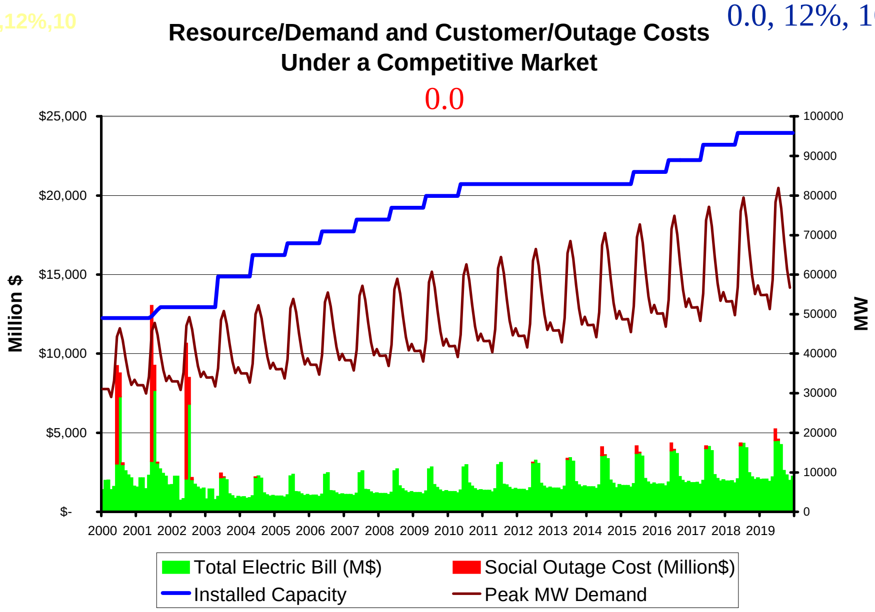
Movie of Demand Elasticities

37

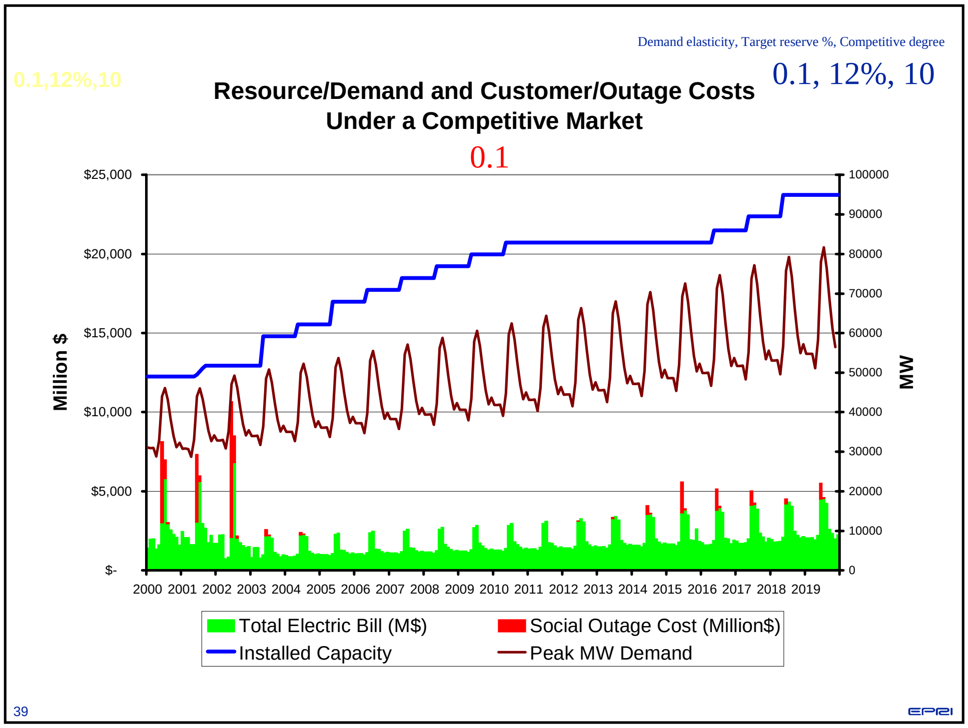
Demand elasticity, Target reserve %, Competitive degree

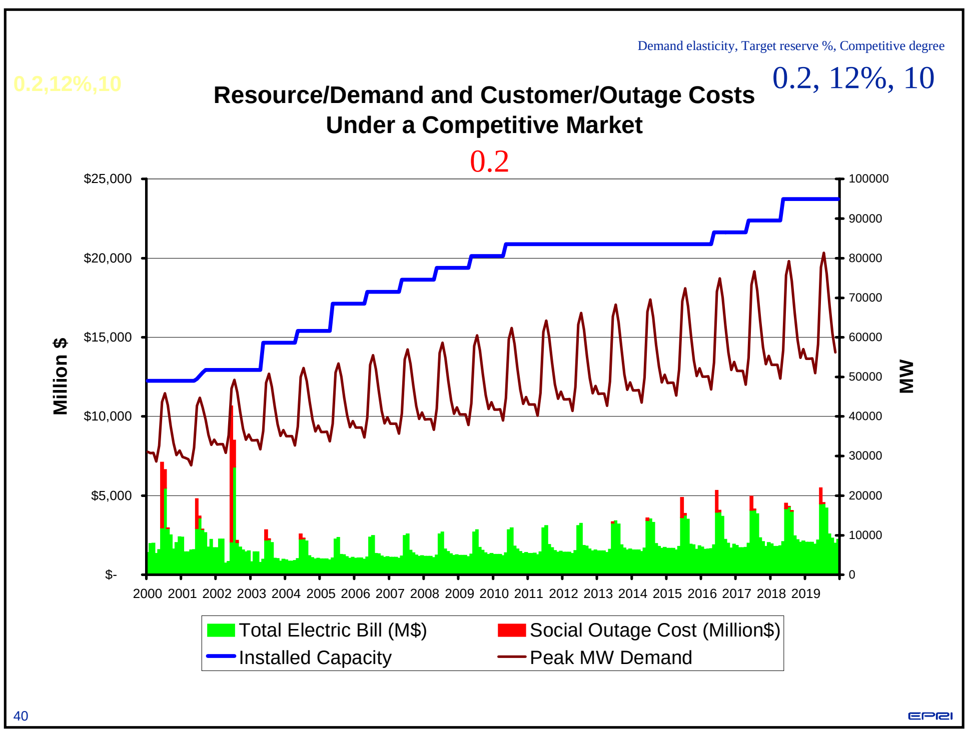
0.0,12%,10

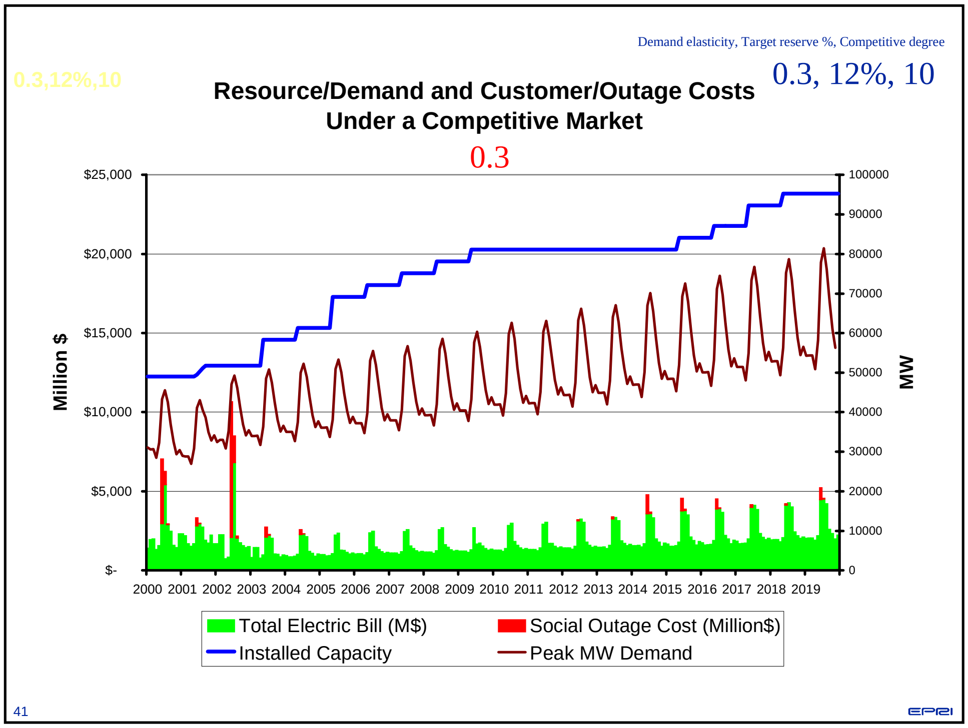
0.0, 12%, 10

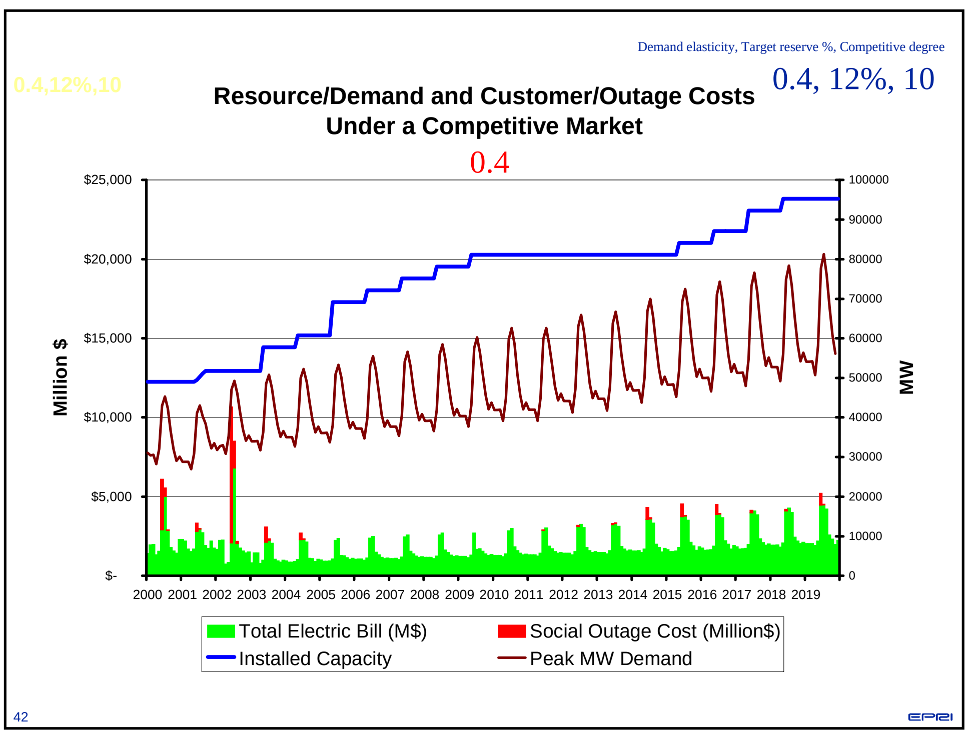


38


Demand elasticity, Target reserve %, Competitive degree
0.1,12%,10
0.1, 12%, 10
Resource/Demand and Customer/Outage Costs
Under a Competitive Market
0.1
$25,000
$20,000
$15,000
$10,000
$5,000
$-
Million $
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
MW
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity
Peak MW Demand
39
EPRI


Demand elasticity, Target reserve %, Competitive degree
0.2,12%,10
0.2, 12%, 10
Resource/Demand and Customer/Outage Costs
Under a Competitive Market
0.2
$25,000
$20,000
$15,000
$10,000
$5,000
$-
Million $
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
MW
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity
Peak MW Demand
40
EPRI


Demand elasticity, Target reserve %, Competitive degree
0.3,12%,10
0.3, 12%, 10
Resource/Demand and Customer/Outage Costs
Under a Competitive Market
0.3
Million $
$25,000
$20,000
$15,000
$10,000
$5,000
$-
MW
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity
Peak MW Demand
41
EPRI

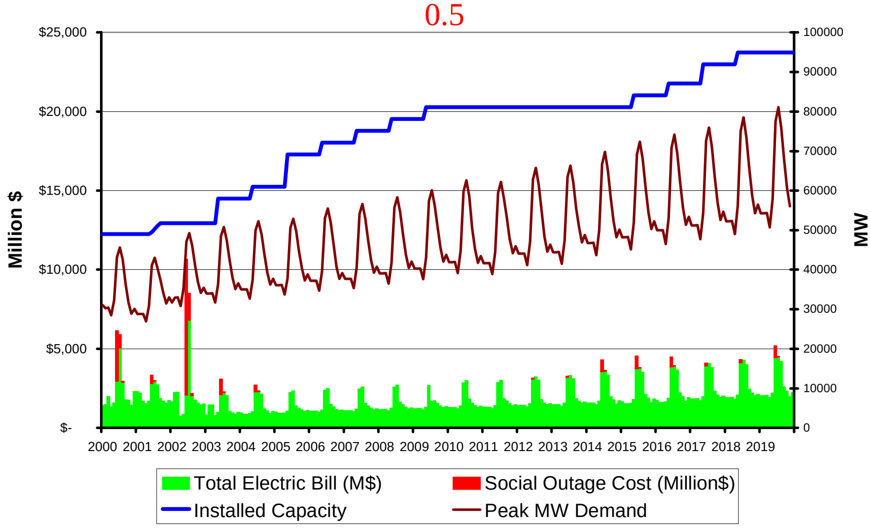

Demand elasticity, Target reserve %, Competitive degree
0.4,12%,10
0.4, 12%, 10
Resource/Demand and Customer/Outage Costs
Under a Competitive Market
0.4
Million $
$25,000
$20,000
$15,000
$10,000
$5,000
$-
MW
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity
Peak MW Demand
42
EPRI

Demand elasticity, Target reserve %, Competitive degree

0.5, 12%, 10

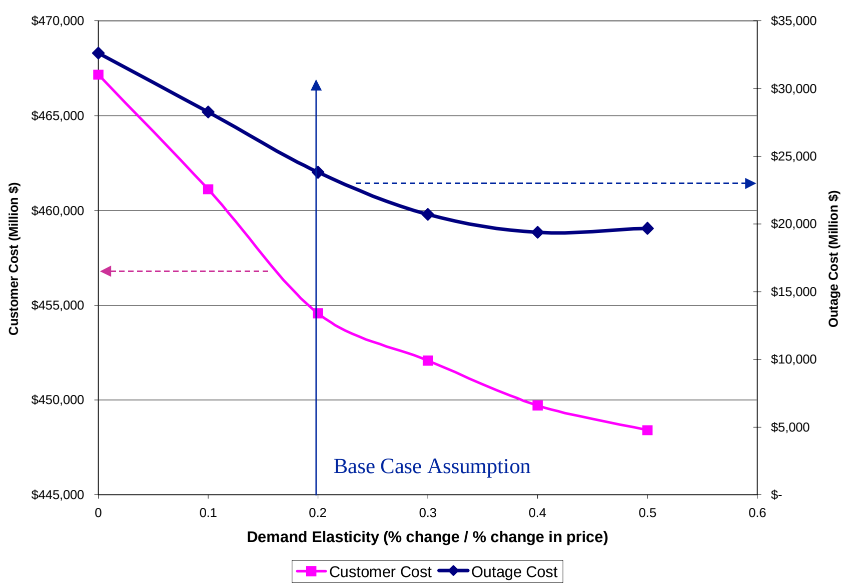
0.5,12%,10

Resource/Demand and Customer/Outage Costs Under a Competitive Market



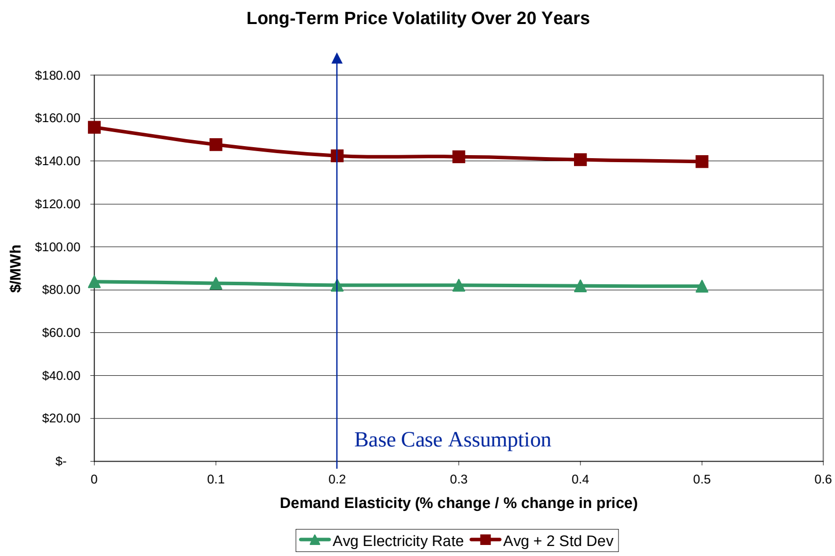
EPRI

Effect of Demand Elasticity on Customer Cost and Outage Cost



EPRI

Effect of Demand Elasticity on Long-term Price Volatility



Movie of Load Growth

Demand elasticity, Target reserve %, Competitive degree

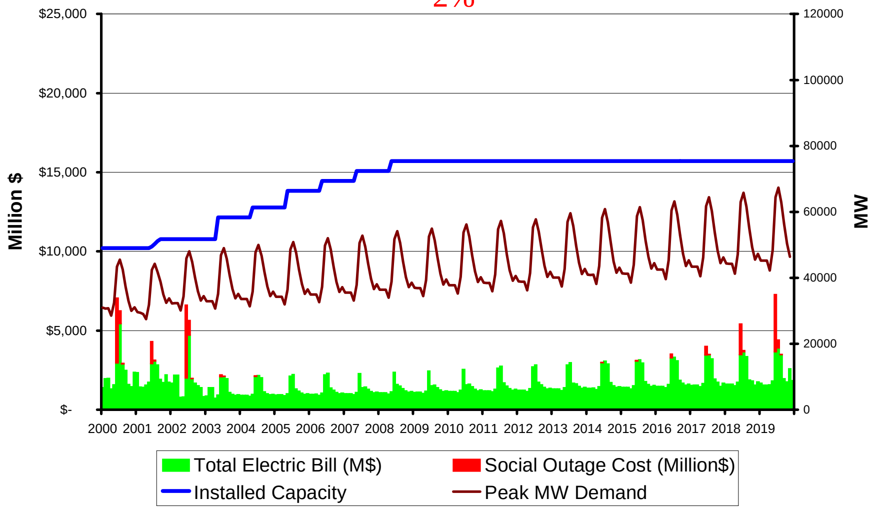
0.2,12%,10,2%

0.2, 12%, 10

Resource/Demand and Customer/Outage Costs Under a Competitive Market

2% load growth

2%



EPRI

Demand elasticity, Target reserve %, Competitive degree

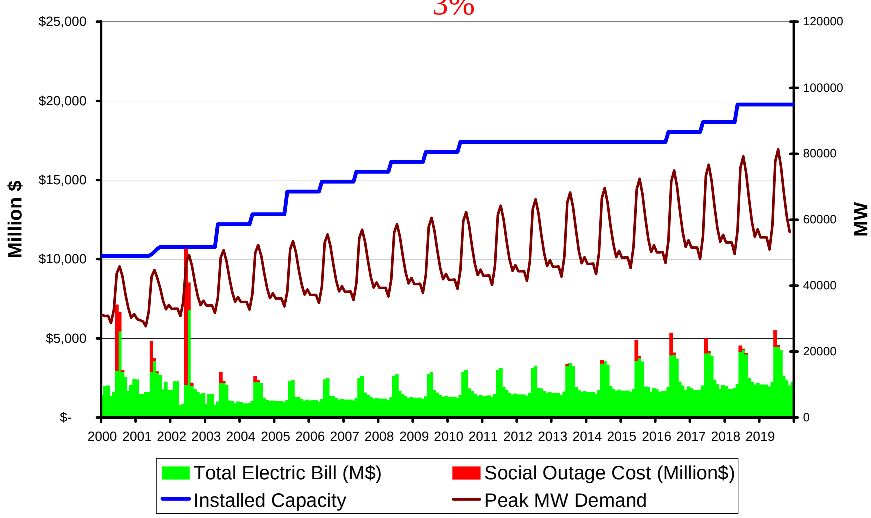
0.2,12%,10,3%

0.2, 12%, 10

Resource/Demand and Customer/Outage Costs Under a Competitive Market

3% load growth

3%



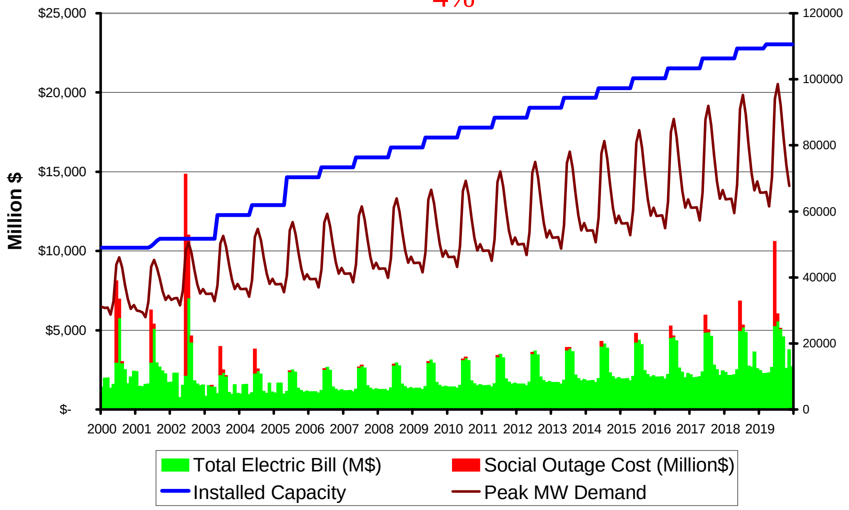
EPRI

Demand elasticity, Target reserve %, Competitive degree

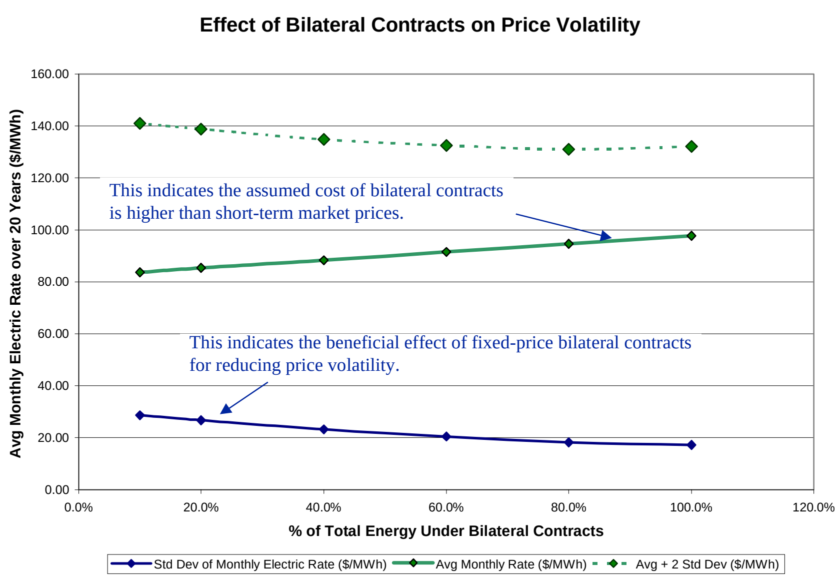
0.2,12%,10,4%

0.2, 12%, 10

4% load growth



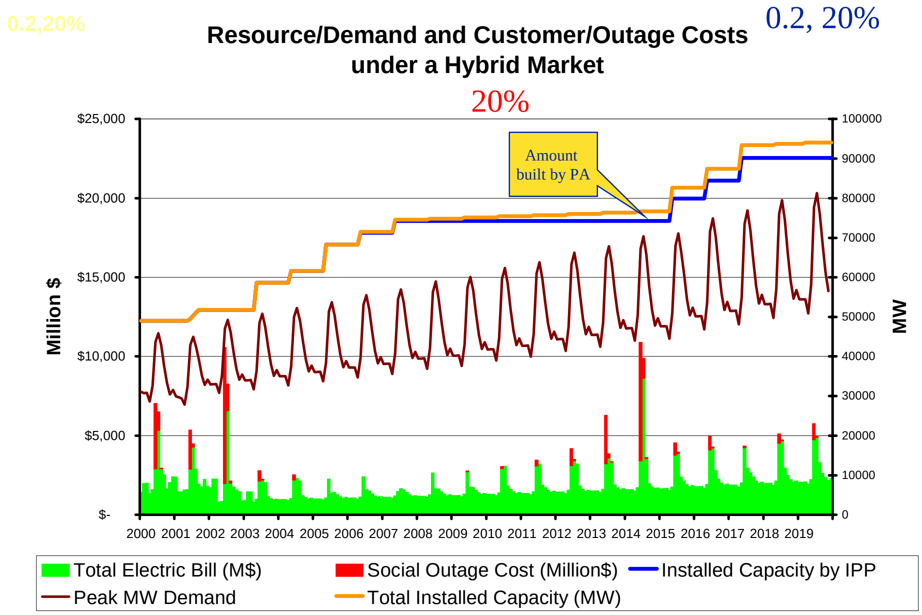
Effect of Bilateral Contracts on Long-term Price Volatility

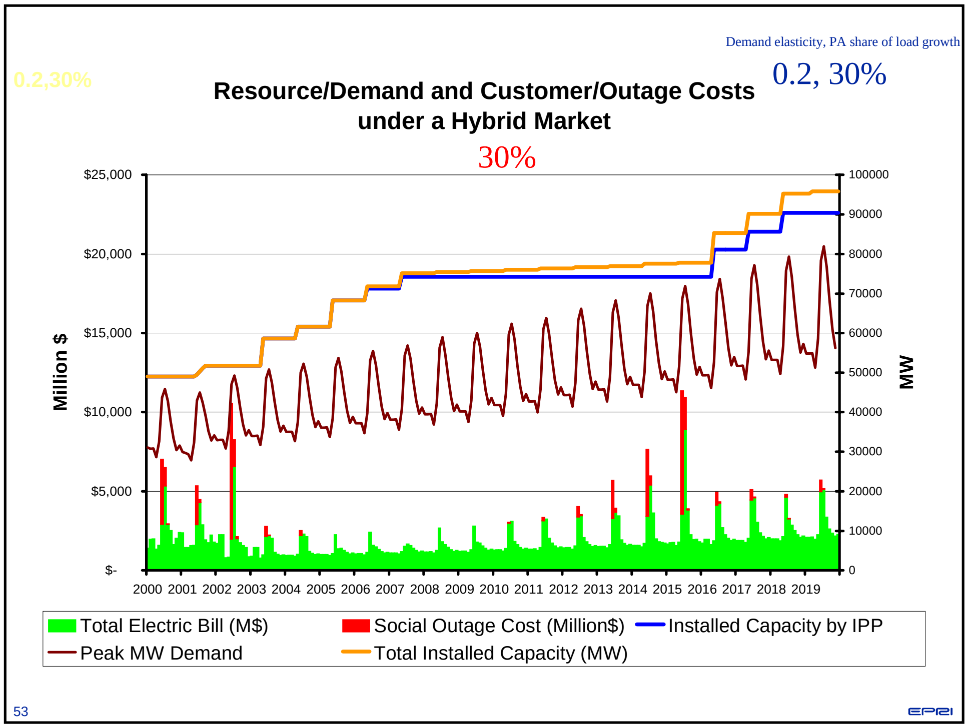


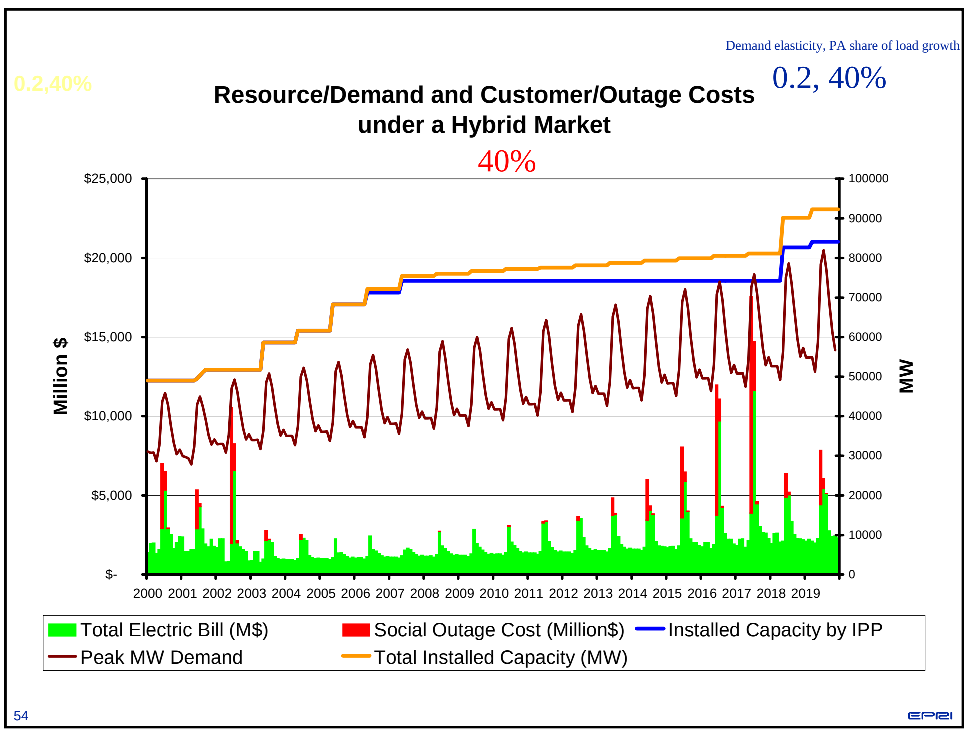
Movie of PA Supply Rule

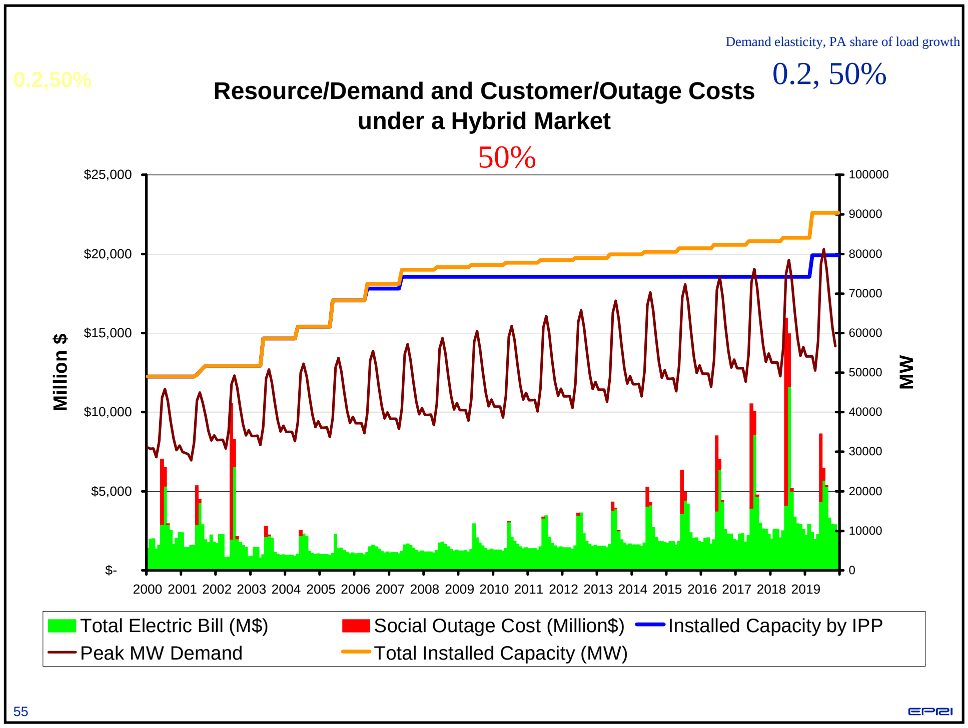
Demand elasticity, PA share of load growth

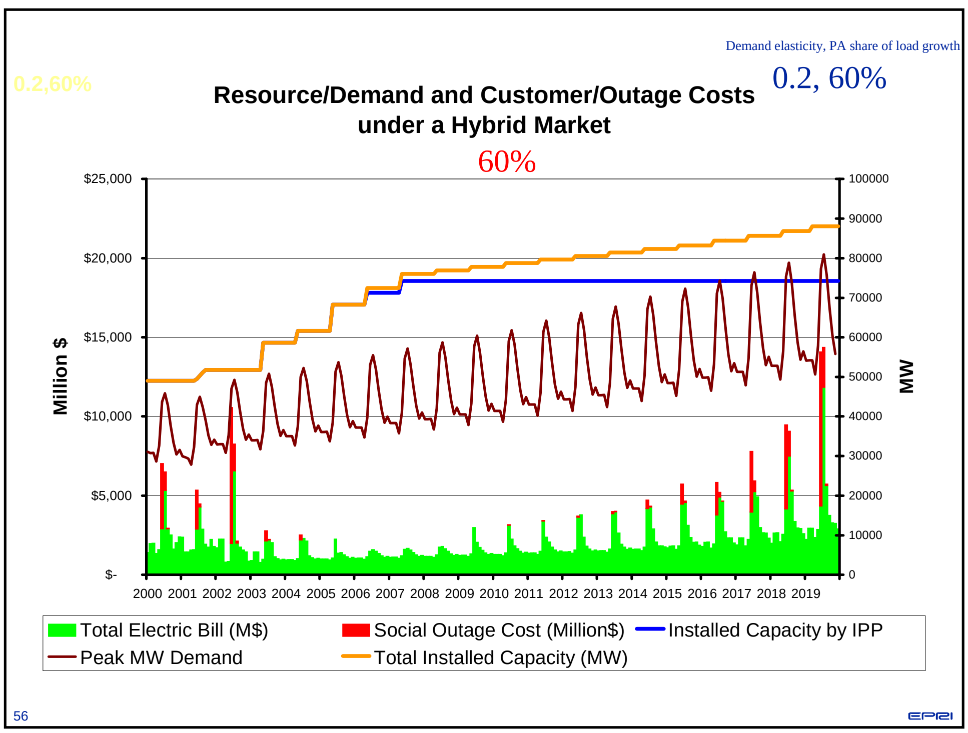
0.2, 20%




Demand elasticity, PA share of load growth
0.2,30%
0.2, 30%
Resource/Demand and Customer/Outage Costs
under a Hybrid Market
30%
Million $
MW
$25,000
$20,000
$15,000
$10,000
$5,000
$-
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity by IPP
Peak MW Demand
Total Installed Capacity (MW)
53


Demand elasticity, PA share of load growth
0.2,40%
0.2, 40%
Resource/Demand and Customer/Outage Costs
under a Hybrid Market
40%
Million $
MW
$25,000
$20,000
$15,000
$10,000
$5,000
$-
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity by IPP
Peak MW Demand
Total Installed Capacity (MW)
54


Demand elasticity, PA share of load growth
0.2,50%
0.2, 50%
Resource/Demand and Customer/Outage Costs
under a Hybrid Market
50%
Million $
$25,000
$20,000
$15,000
$10,000
$5,000
$-
MW
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity by IPP
Peak MW Demand
Total Installed Capacity (MW)
55
EPRI


Demand elasticity, PA share of load growth
0.2,60%
0.2, 60%
Resource/Demand and Customer/Outage Costs
under a Hybrid Market
60%
Million $
$25,000
$20,000
$15,000
$10,000
$5,000
$-
MW
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
Total Electric Bill (M$)
Social Outage Cost (Million$)
Installed Capacity by IPP
Peak MW Demand
Total Installed Capacity (MW)
56
EPRI

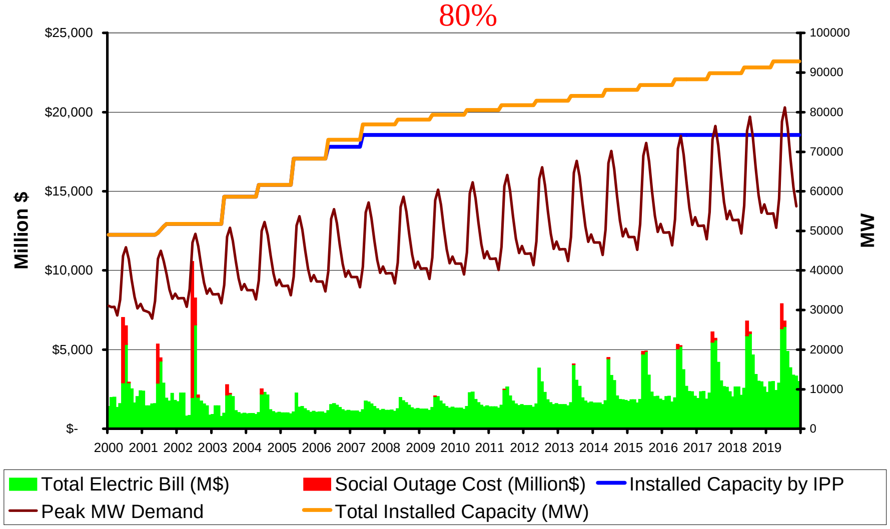
Demand elasticity, PA share of load growth

0.2,80%

0.2, 80%

Resource/Demand and Customer/Outage Costs under a Hybrid Market

80%



EPRI

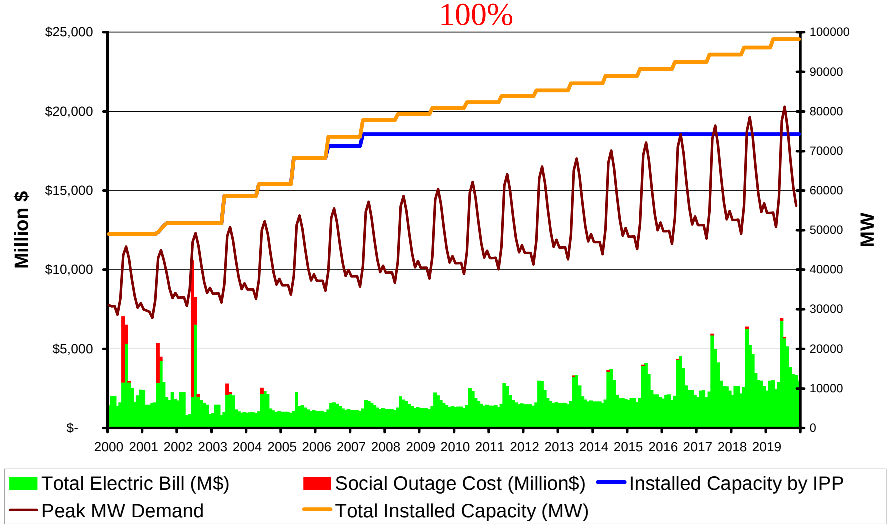
Demand elasticity, PA share of load growth

0.2,100%

0.2, 100%

Resource/Demand and Customer/Outage Costs under a Hybrid Market

100%

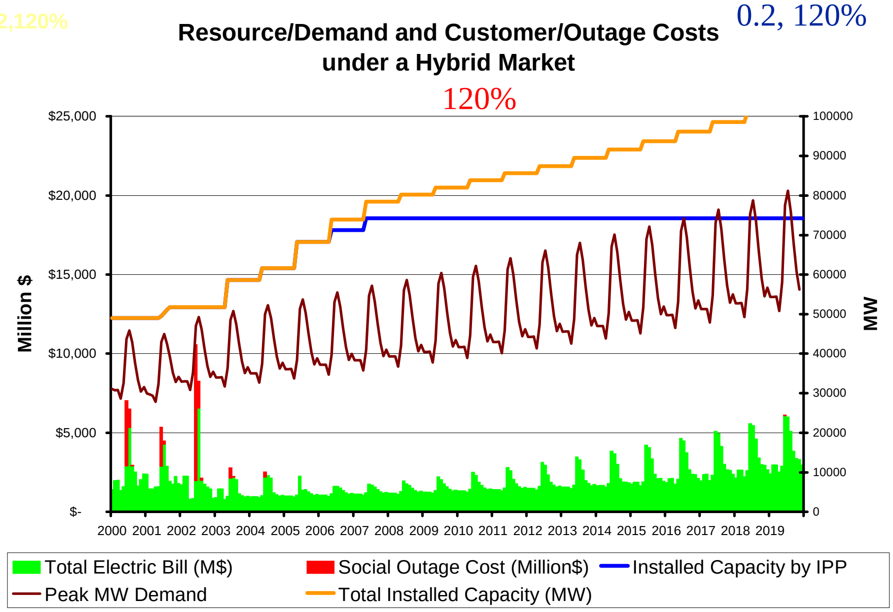


EPRI

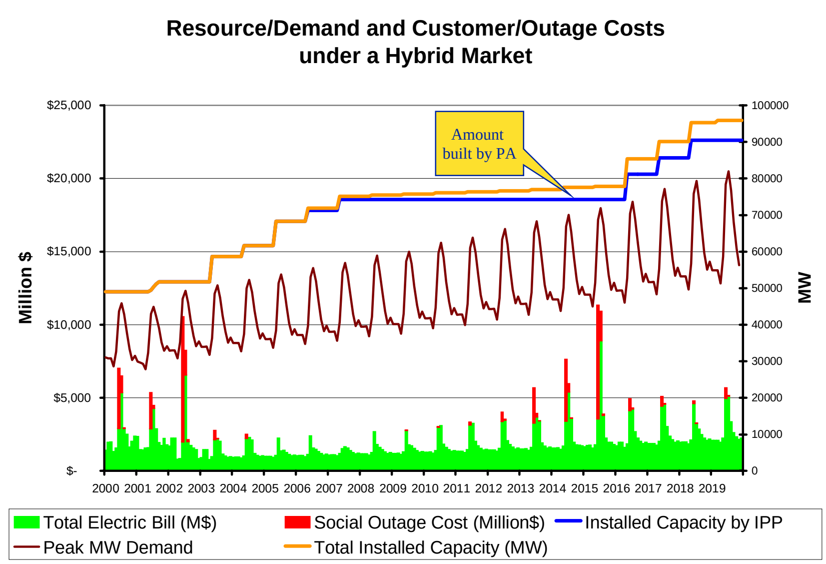
Demand elasticity, PA share of load growth

0.2,120%

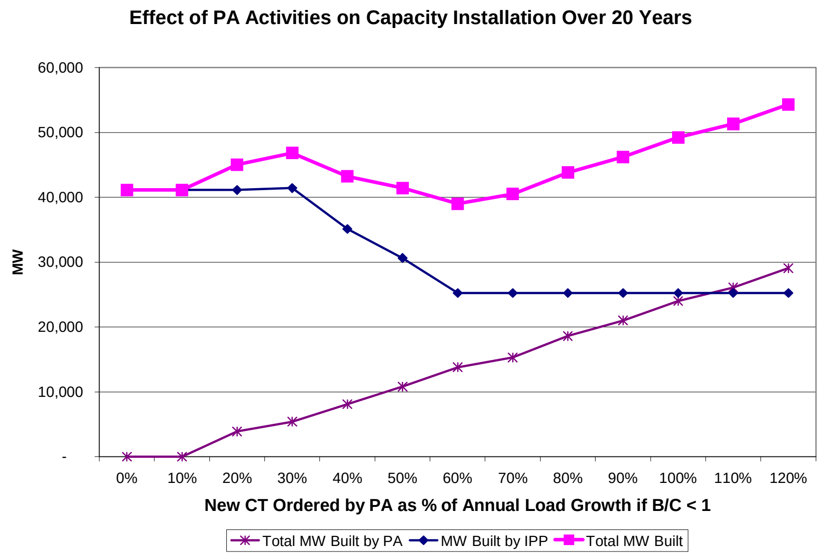
0.2, 120%



A Hybrid Market Scenario PA Supply Rule is 30% of Load Growth

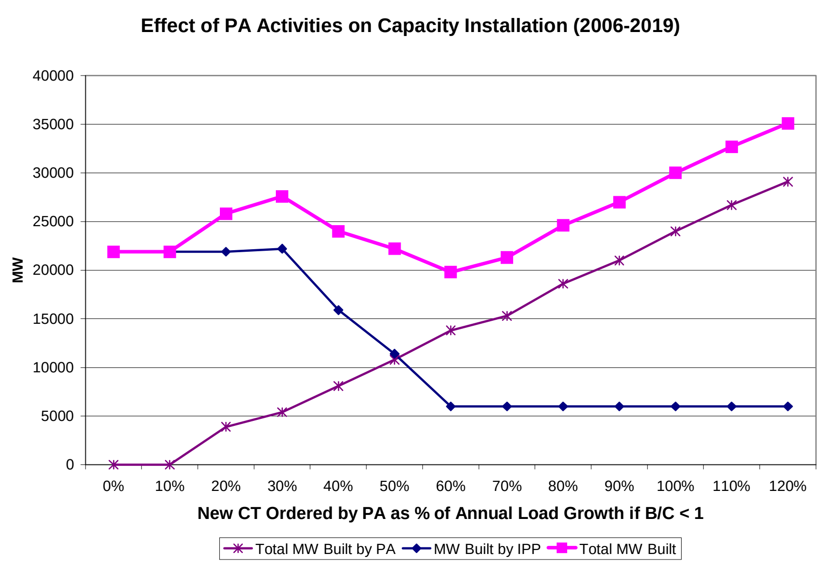


Effect of PA's Supply Rule on Total Capacity Installed over 20 Yrs



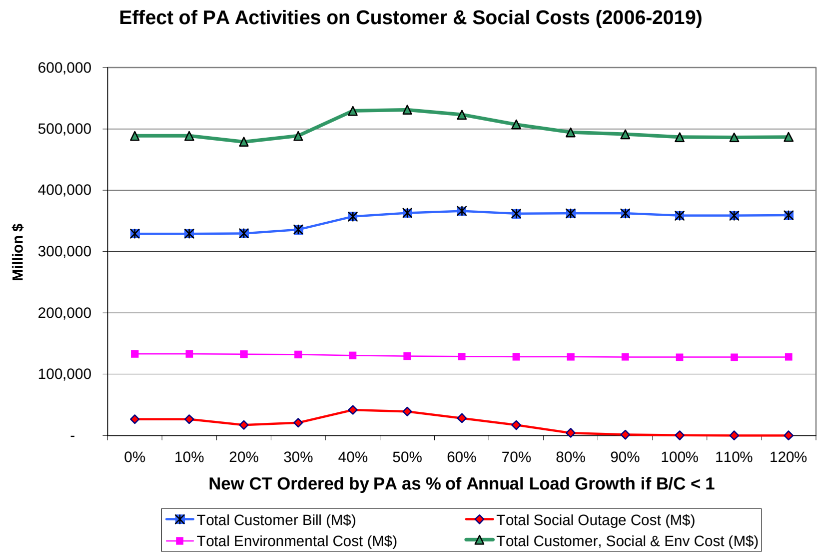
61

Effect of PA's Supply Rule on Total Capacity Installed 2006-19



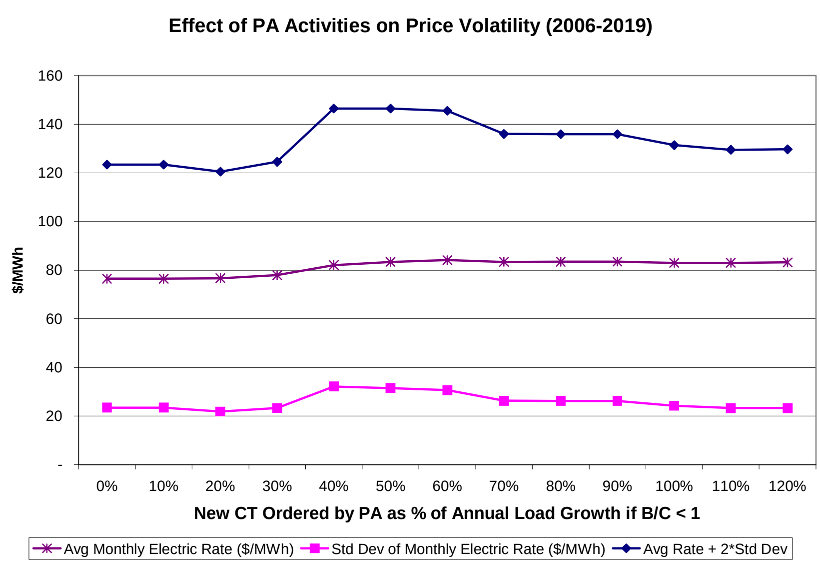
62

Effect of PA's Supply Rule on Customer & Social Costs 2006-19



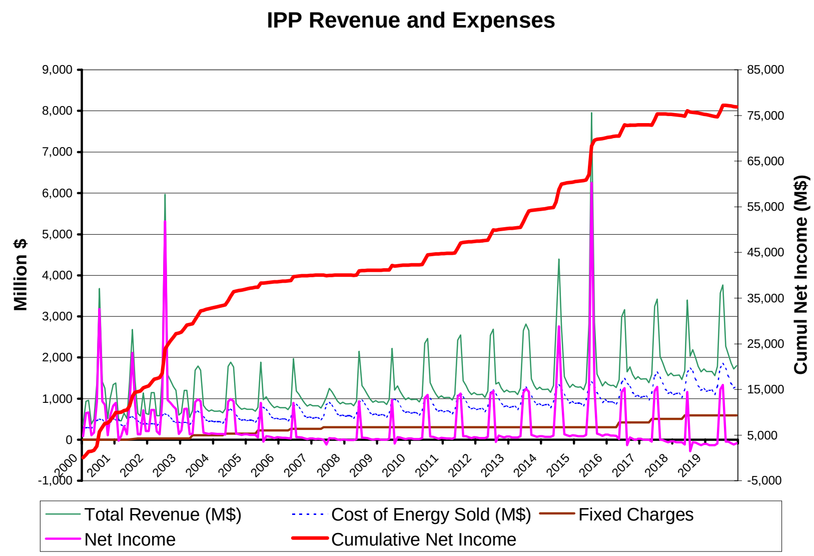
63 EPRI

Effect of PA's Supply Rule on Price Volatility (2006-2019)



64 EPRI

IPP Cash Flows in Hybrid Market If PA's Supply Rule is 30%



Summary

- Desirable features of a healthy and robust power market:
  - Efficient
  - Reliable
  - Ability to withstand shocks
- Is a competitive power market naturally healthy and robust in the long term?
- If not, what improvements should be made?

Panel Session 1

Long Term Objectives of the California Wholesale Power Market

CEC Workshop on Alternative Wholesale Market Structures

Eric Charles Woychik
Strategy Integration Inc.
estrategy@mindspring.com

Outline of Comments

- 1. Long Term Objectives for CA
- 2. Lessons Learned from CA and Others
- 3. Other Wholesale Markets
- 4. Structural Shortcomings of purely competitive markets

1. Long-Term Objectives

- Minimum consumer costs
- Efficient operations
- Maximum use of assets
- Optimal investment
- Adequate risk management
- Maximum benefit for society

Objectives Vis-a-Vis End State?

- What end-state is acceptible?
- Threshold decisions:
  -- Market structure
  -- Governance of grid operation & markets
- We need objectives that are specific to market structure and governance

Objectives for Competition

- Lower costs and prices
- Provide better incentives for operation and consumption
- Improve regional trading
- Maximize asset use
- Properly stimulate new innovation & investment

2. What Lessons Learned?

- "It's the market structure stupid"
- Price signals do matter; average pricing that fails to reflect grid physics does not work
- Price volatility should drive demand response and locational price mitigation
- Thus, "market pricing" with zones, without real locational differences doesn't work

California Lessons Learned: I

- Acted too fast without full assessment
- Market participants find market flaws & exploit them – the incentives are huge
- Governance by market participants, particularly if grid and market rules fail – a disaster
- Costs of flaws and failure: now $50 billion
- Marginal cost bids resolve major problem & provide partial-competition

California Lessons Learned: II

- Market power abuse: high irrational prices for all services – flawed market structure
- Biased governance of market structure
- Inadequate bilateral contracts to hedge spot prices
- Retail prices that did not reflect wholesale prices and lack of price-responsive demand

3. Other Competitive Markets

- Back-yard problem is WSCC -- the other "competitive" market (outside CA)
- CA cannot solve its problem and ignore the rest of WSCC, nor can WSCC ignore CA
- CA "tail" wagged the WSCC dog, to death
- Northern Europe; integrated of necessity
- North East (U.S.); is integrating of necessity

Lessons Learned – All Data Points

- Chile/Northern Europe
- England/Wales & Scotland
- New Zealand
- Australia
- Pennsylvania, Jersey, Maryland (PJM)
- New England
- New York
- Alberta
- Texas
- Interconnected markets
- California

Lessons Learned – Summary: I

- Chile, Northern Europe, and PJM: proper market structure, then market works quite well
- England/Wales & Scotland: poor market structure means greater regulation & costs
- Early New Zealand: model failed to reflect grid operation, so provided poor incentives
- Australia: some market power abuse by generators
- Early PJM -- poor governance & poor incentives

Lessons Learned – Summary: II

- New England: market structure/generation control
- New York: software & settlements
- Alberta: market structure & transition
- Texas: some market power abuse; market structure
- Interconnected markets must use same model or fully address seams issues (not yet successful)
- California – failed market structure, market power abuse, and discriminatory governance

Three Lessons for All Markets

- 1st: Market Structure – get the incentives right
- 2nd: Independent Governance of Grid & Market Operations
- 3rd: Organize to meet Regional market needs

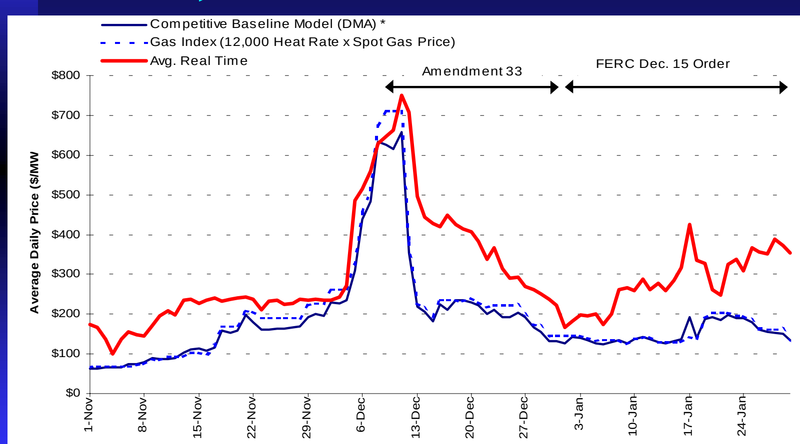
4. Structural Shortcomings of Competitive Markets

- Local market power if single or few generators are needed for voltage/stability
- Tight-supply-demand creates opportunity for physical or economic withholding
- Complex, sequential markets are a recipe for market power abuse
- Market structure must be KISS, one-shot, physical-market in synch, automatic PBCs

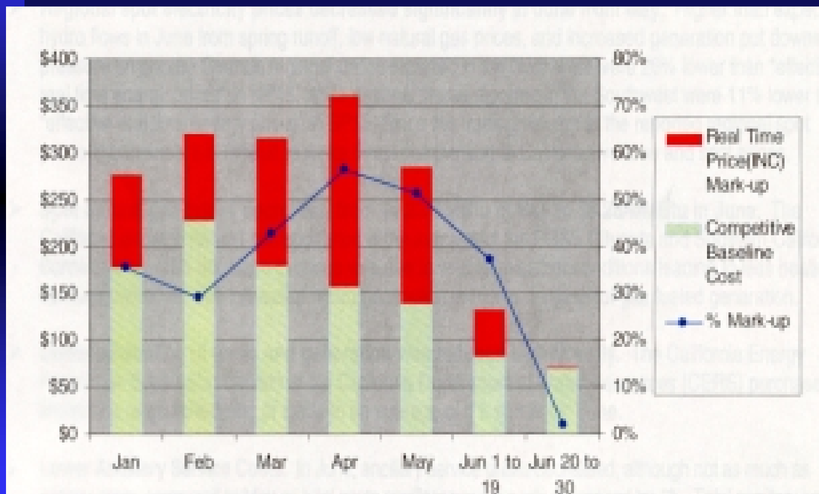
Recurring Problem: Market Abuse by Manipulation of Bids

- Generators bid to artificially raise prices
- Little consumer response to prices
- Need rules to create "honest bidding"
- Excessive bids are "economic withholding" of generation -- the major problem
- Plant bid caps address this problem directly, when workable competition does not exist

Baseline Competition, Real-Time Prices, and Gas Prices



Price-Cost Mark-up in RT: '01



How to Address Shortcomings?

- FERC wants competitive markets, "best practices," and RTO business plans
- Single, best practices market structure and governance are exhibited in PJM
- PJM + NEISO + NYISO is *fait accompli*
- What about West-wide RTO?
- Stuck in limbo-land; politics and lack of leadership

Long-Term Objectives of the California Wholesale Power Market

LAWRENCE LINGBLOOM

Senate Energy, Utilities and Communications Committee

California Energy Commission Workshop
November 7, 2001

Q1: What are the long-term objectives of the wholesale market?

- For the vast majority of the public, the objective is the same as it's always been - facilitate reliable and reasonably-priced retail service.
- The market's structure is simply a means to achieve this end (or not) - if the end is achieved, the means will be justified in the minds of consumers (and vice-versa).
- Regulated or unregulated, public or private, spot market or long-term contracts - these details are incidental to the fundamental issues - Do the lights come on when I flip the switch and what's the number when I open the bill?
- In California, we'll try to integrate certain policy objectives into whatever the scheme is - renewable and clean generation, energy efficiency, conservation, etc.
- Most consumers don't want to have to care about the wholesale market - a successful wholesale market should be invisible to those who want to ignore it.

Senate Energy, Utilities and Communications Committee

The public is in no mood to take risks.

The Energy Crisis, as well as other recent economic and political events, has underscored public sentiment for reliable, stable electricity service.

Key factors:

- Fundamentals that precipitated the Crisis - supply/demand, immaturity of the market, rampant opportunism among market participants.
- Failure of FERC to provide a timely response.
- Condition of economy in general, and market performance of energy companies in particular.
- Security threats.

Senate Energy, Utilities and Communications Committee

Q2: What lessons have we learned about the importance of market structures?

Lesson #1 - If you're going to have a market, it needs some structure.

- Optimism and a four-year transition period do not a market make - over-confidence in the promises of restructuring caused nearly everyone involved to downplay or ignore fundamental deficiencies.
  - Lack of "demand elasticity"
  - Mismatch between capacity needs and investor motivation
- There's no such thing as a "perfect" market - markets need discipline in the form of meaningful monitoring and sanction authority.
- Market participants with private capital on the line are clever and ruthlessly opportunistic. If they don't succeed in creating a market structure that advantages them, they'll either find a way around it, or abandon it. Neither outcome is conducive to producing enduring consumer benefits.

Senate Energy, Utilities and Communications Committee

Q3: Are other markets functionally competitive, or have they simply benefited from surplus capacity?

- Don't really know if eastern markets are "functionally competitive" - not even sure if I know what that means.
- Experts suggest that they have better supply conditions, or at least more discipline, and more affirmative market power mitigation measures than California.
- These markets also seem to benefit from fewer moving parts, and may be easier to manage.
- Difficult to isolate competitiveness from adequate, or surplus, capacity - adequate capacity is one of the conditions needed for competition, but does not necessarily lead to a competitive result.
- These markets seemed to have outperformed California in that they've produced a better bottom-line result for consumers.
- Can't be sure of the reasons, or whether it's sustainable.

Senate Energy, Utilities and Communications Committee

Q4: What are the structural shortcomings of a purely competitive wholesale market?

(What are the *competitive* shortcomings of a purely *structural* wholesale market?)

- A "purely competitive" (unregulated, private capital investments) electricity supply system is inconsistent with state policy objectives.
- For one thing, it exposes consumers of an essential commodity to both the boom and bust tendencies of the energy business. This means unpredictable prices and unpredictable service - both of which are unacceptable.
- A purely competitive system is also unlikely to adequately reflect the full range of public values related to electricity service
  - Air quality
  - Land use
  - Social justice

Senate Energy, Utilities and Communications Committee

What market structure will yield reliable service, reasonable prices and a return sufficient to attract and sustain committed participants?

- I don't know - I'll leave it to the experts to speculate, although I'm not convinced that any of them knows with any certainty.
- Unfortunately, the concept of competition in electricity markets has been contaminated by folks pursuing a business model that is fundamentally anti-consumer and unsustainable.
- At this point, it seems that virtually all market participants have mistreated one another, and mistreated consumers.
- Significant changes in structure and attitude are needed to restore a functional market.
- Legacy of the Energy Crisis is that there's no longer much room for experiment.

Senate Energy, Utilities and Communications Committee

California Consumer Power and Conservation Financing Authority

California Energy Commission/EPRI
Market Design Workshop
November 7, 2001


Kellan Fluckiger
Senior Advisor to the Chair and CEO

California Consumer Power and Conservation Financing Authority

<u>OLD DESIGN</u>

- Obligation to Serve
- Rate Base Paradigm
- Regulated Rate of Return
- Public Policy Intervention
- Unwritten Policy Backdrop

California Consumer Power and Conservation Financing Authority

<u>OLD DESIGN</u>

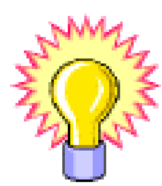
<u>Obligation to Serve</u>

- Defined Area
- Limited Competition
- Integrated Resource & Transmission Planning
- Single Use for Generation and Transmission Assets
- Obligation to Build

California Consumer Power and Conservation Financing Authority

<u>OLD DESIGN</u>

<u>Rate Base Paradigm</u>



- Build as much as needed
- Over building is OK
- Safe & Stable
- Cost Plus leads to high rates

California Consumer Power and Conservation Financing Authority

OLD DESIGN

Regulated Rate of Return

- Reliance on 3rd Party Review
- Prudence Audits
- Expensive Oversight
- How much is enough?

California Consumer Power and Conservation Financing Authority

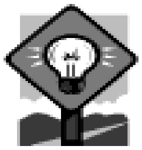
OLD DESIGN

Public Policy Intervention

- Clean Air
- Alternative Generation Technologies
- Efficiency Improvements
- Conservation
- Environmental Impact Mitigation

California Consumer Power and Conservation Financing Authority

OLD DESIGN



Unwritten Policy Backdrop

- Everyone is entitled to electricity
- Everyone is entitled to a reasonable price

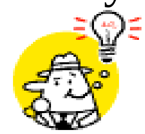
California Consumer Power and Conservation Financing Authority

CURRENT DISASTER

- No Obligation to Serve or Build
- No Integrated Resource Planning
- Full Reliance on Spot Market
- Questionable Market Behavior
- No Transitional Period
- Inadequate Price Signals

California Consumer Power and Conservation Financing Authority

CURRENT DISASTER



No Obligation to Serve or Build

- Utilities relieved of historic obligation to serve
- Provider of last resort became the ISO
- Stranded cost recovery incented utilities to rely on ISO
- Utilities out of generation building business
- Complete reliance on markets to build additional plant

California Consumer Power and Conservation Financing Authority

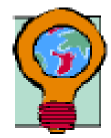
CURRENT DISASTER

No Integrated Resource Planning

- Transmission and generation no longer planned together
- Transmission system never designed to support a market
- No entity had the obligation to ensure sufficient capacity
- Least cost analysis turned over to market
- No allowance for long lead times

California Consumer Power and Conservation Financing Authority

CURRENT DISASTER



Full Reliance on Spot Market

- Utilities divested large quantities of generation
- Utilities did not recreate a supply portfolio
- Spot prices do not reflect long run marginal costs
- Market design incented under scheduling

California Consumer Power and Conservation Financing Authority

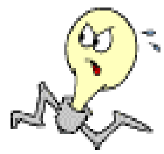
CURRENT DISASTER

Questionable Market Behavior

- Bids do not reflect marginal cost
- Historic outage rates
- Radically different incentives to new generation owners
- Questions of physical and economic withholding
- Market design incents gaming

California Consumer Power and Conservation Financing Authority

CURRENT DISASTER

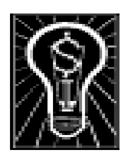


No Transitional Period

- Lack of vesting contracts
- Lack of price to beat
- No preparation of consumers
- No preparation for end of stranded cost recovery period

California Consumer Power and Conservation Financing Authority

CURRENT DISASTER



Inadequate Price Signals

- No consumer education
- No preparation for rate shock
- High volatility provides poor signal to business interests
- Transmission signals completely ineffective
- Market dysfunctionality destroyed what price signals there were

California Consumer Power and Conservation Financing Authority

<u>FUTURE DESIGN</u>

- Capacity Requirement
- Limited Day Ahead Market
- Obligation to Perform
- Peak Management
- Full Regional Coordination

California Consumer Power and Conservation Financing Authority

<u>FUTURE DESIGN</u>

<u>Capacity Requirement</u>

- Load must have capacity requirement
- Load must have reserve requirement
- Phase in period required
- Public power may play a role
- Capacity must be sufficient to ensure competition

California Consumer Power and Conservation Financing Authority

FUTURE DESIGN

Limited Day Ahead Market

- Load is managed as a portfolio
- ISO managed day-ahead market for energy and ancillary services necessary for unit commitment
- Balanced requirement irrelevant at day-ahead stage
- Transmission congestion must be managed day-ahead-feasible schedules
- Size of day ahead market may be limited

California Consumer Power and Conservation Financing Authority

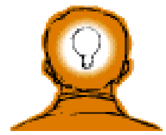
FUTURE DESIGN

Obligation to Perform

- Generators must perform as scheduled
- Penalties for non-performance
- Delivery risk should be re-evaluated

California Consumer Power and Conservation Financing Authority

FUTURE DESIGN

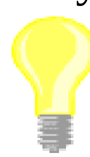


Peak Management

- Competition for peaked capacity is complicated
- Price spiked deemed unacceptable
- Public participation may be required
- Public financing is more economical

California Consumer Power and Conservation Financing Authority

FUTURE DESIGN



Full Regional Coordination

- Single grid management process must be achieved
- Transmission service must truly be open
- Long term transmission products must be available
- Regional timeline coordination essential
- Market design must integrate with other regional processes

Panel Session 2

Means of Reducing Price Volatility in a Competitive Power Market from the Supply and the Demand Sides

What causes price variation?

Carl Blumstein
Nov 7, 2001

What causes price variation?

- Variation in short run costs
- Capacity constraints
  - –physical limits
  - –market power (withholding or, equivalently, high bidding)
- Variation increased by inelastic demand

What are some things we can we do about price variation caused by capacity constraints?

- Shift the supply curve to the right
  - build publicly owned peaking capacity
  - make side payments to increase capacity
  - mitigate market power
- Shift demand curve to the left
  - energy efficiency
- Tilt the demand curve (make demand elastic)
  - real-time pricing
- Hedging contracts for risk averse consumers

What are some potential problems with these strategies?

- Capacity payments either give the wrong price signal or do not eliminate the price variation caused by capacity constraints
- Public investment will drive out private investment (no investment in peakers, under investment in base load plant)
- Uncertainty about price elasticity
- Institutional barriers

Notes from:

Power System Economics

(Structure, Profits, Investment & Caps)

CEC Talk, Nov. 7, 2001

S t e v e n   S t o f t

www . s t o f t . com

The Two Demand-Side Flaws

- Demand-side Flaw 1:
  Lack of Metering and Real-Time Billing
- Demand-side Flaw 2:
  Lack of Real-Time Control of Power Flow to Specific Customers

(See Section 1-1.5)

2

Flaws 1 & 2 Imply Market Effectively Capped

- **The market will not always clear.**
- **When it does not, Flaw 2 forces the SO to set a price.**
- **The limit on how high a price the SO is willing to set will cap the market.**

  - If the SO sets a $ Billion limit and no other law sets a lower limit, the market will have no long-run equilibrium and will periodically self destruct.
  - Every power market has a limit, and in the U.S. they are lower than the theoretical VOLL limit in Australia.

How the SO's Limit Caps All Energy Markets

- **If the SO pays no more than X,**
  it cannot charge more than X.   (say X = $1000 / MWh)
- **There is no reason for consumers to pay more**
  when you can get the best possible power by just taking it.
- **The SO will only charge them $1000 / MWh.**

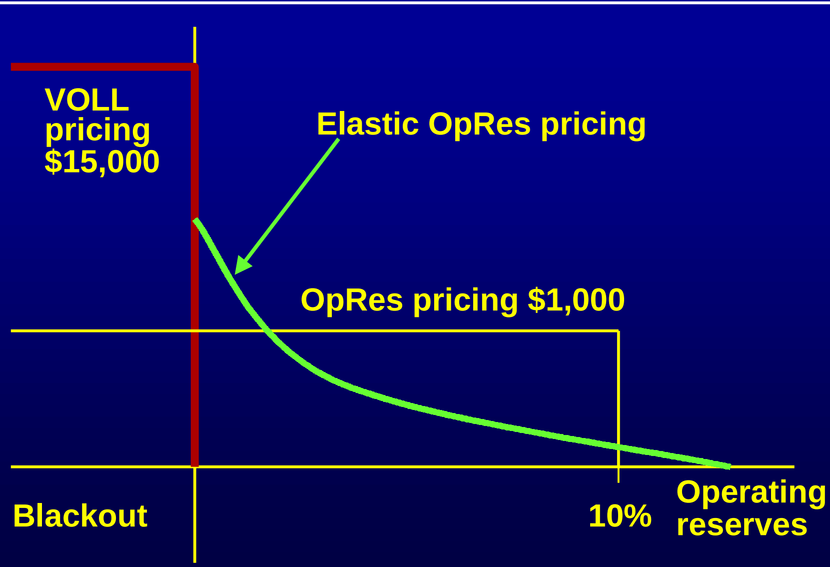
Fact 1:

System operators always have a purchase-price limit.

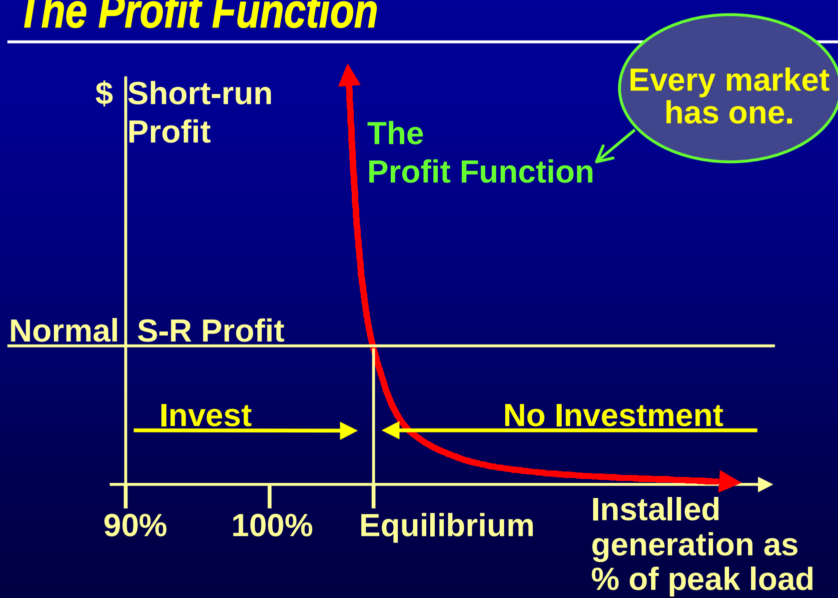
Fact 2:

The real-time SO price limit caps the entire market.

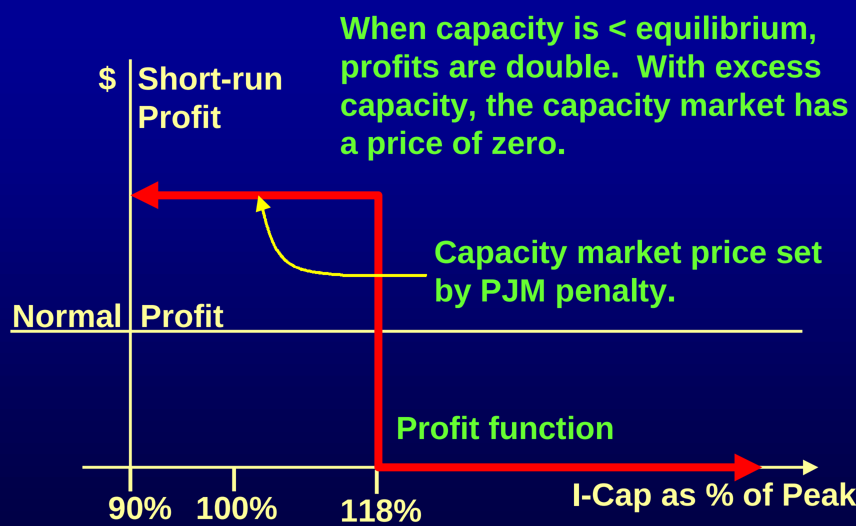
Some Possible Price Limits



The Profit Function



Installed-Capacity Approach



The Profit Function (#2)

- **The profit function controls**
  - Investment.
  - Reliability (blackouts).
  - The success or failure of electricity deregulation.
- **The profit function can be and is controlled by regulators**
  - Some of these are disguised as NERC engineers.
- **What controls the profit function?**
  - Operating reserve requirements.
  - How much (X) the SO pays when short of reserves.
  - Installed capacity requirements.

How to Have Too Much Reliability

You don't need extreme prices!

Method 1:
- Require 40% operating reserves.
- Set the price cap, X, to $300/MWh
- Have no installed generation capacity requirement.

Method 2:
- Require 10% operating reserves (traditional level).
- Set the price cap, X, to $30,000/MWh
  (Pay this whenever reserves get below 10%).
- No capacity requirement.

Method 3:
- Require 2% operating reserves
- Price cap, X = $300
- Capacity requirement set to 150% of peak load.

How to Have too little Reliability

Method 1:
- Operating reserve requirement at 2%.
- Price cap at $1000/MWh.
- No capacity requirement.

Method 2:
- Operating reserve requirement at 10%.
- Price cap at $80/MWh.
- No capacity requirement.

The Right Profit Function

- Should give correct level of reliability.
- Should avoid boom-bust investment cycles.
- Market fundamentalists believe a free power market does #1 automatically. (Not yet.)
- Point #2 has been generally overlooked, but Chairman Hoecker recognized it in his final message on California.

> ***"Other regions address this with a reserve obligation supported by a capacity market. Without it, says one analysis, 'the market is designed to produce periodic reliability crises with energy price booms followed by price busts.' "***
>
> *Hoecker quoting CERA, "A Crisis By Design: California's Electric Power Crunch," September 2000, at 2.*

The Right Profit Function (#2)

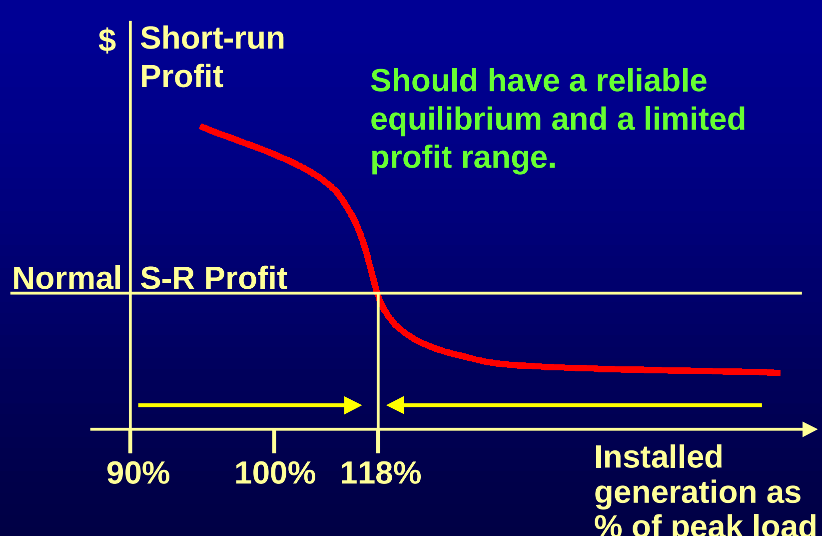
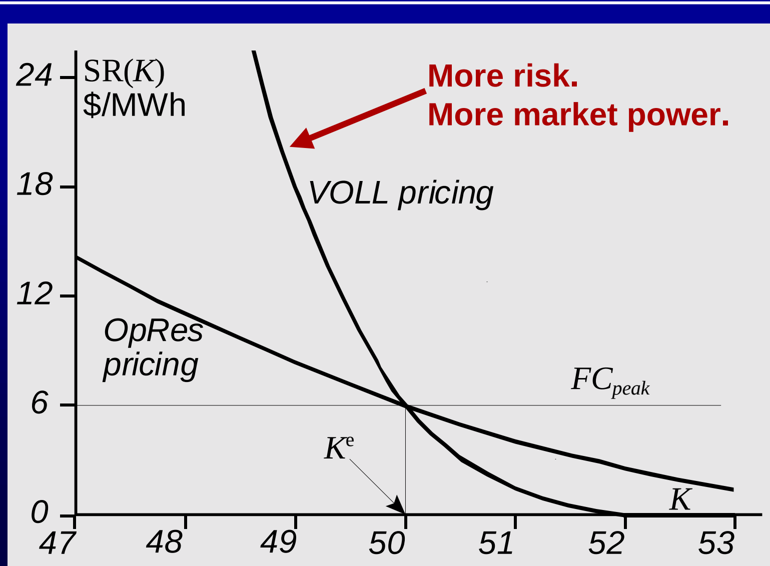


Figure 2-7.1 Profit Functions



Conclusion on Profit-Function Design

- Without realizing it, regulators have been designing our price spikes.
- Regulators are just starting to notice the profit function.
- NERC, FERC, ISO-Transco designers, & system operators need to cooperate and coordinate.
- Until these market fundamentals are understood we can expect more instability in power markets.

The 2nd Biggest Problem

- Competition between System Operators
- The low-spike operator loses reserves to the high-spike operators.
- They are forced to compete over how high their spikes are.
- They don't understand they are cutting their own throats.
- They have a "pay anything for reliability" mentality.
- "Out of market" they do.
- Only FERC / NERC can bring about the necessary cooperation.

No Matter What California Decides to Do, It Should At Least . . .

Presentation to the California Energy Commission

John D. Chandley

November 7, 2001

The central question for California: Should it reform the Market? Or abandon the market?

Reforming the market may appear very hard, but we know what to do.

- **We have examples of successful markets**
- **A standardize market design is emerging, and it works**
- **FERC supports this, and they've learned a lot (but is it enough?)**
- **We generally know what a supporting state regulatory structure needs to do to make the market workable**

Abandoning the market is also very hard, and it's not clear how you go back to a fully regulated system. What does it mean?

- **Can/should we reassemble the IOUs as vertically integrated?**
- **Do you seize existing IPPs? Ban new IPPs?**
- **Would you throw away regional dispatch?**
- **Disconnect from the West, or defy FERC?**

There are actions that make sense, no matter which choice California makes

1. Preserve the regional system operator.

   - **A pooled dispatch for the 3 IOU regions allows efficient cost savings. Separate dispatches are likely to be less efficient.**
   - **To capture these benefits, you need to return to the familiar principles of a security-constrained, economic dispatch.**
   - **Your system operators know how to do this, but rules prevent it.**

2. Keep using offers/bids to define an efficient dispatch.

   - **Bids get the information the dispatcher needs from generators that aren't affiliated or in the same utility, creating a level playing field.**
   - **Carefully design bidding rules can avoid the problems you had.**
   - **Bidding processes provide a logical mechanism to organize "spot" or "economy energy" purchases and sales.**

Other actions that make sense, no matter what California decides to do about markets

3. Keep the regional dispatch process open to all parties. There are beneficial "spot" trades that can be easily made through a bid-based economic dispatch process:

   - **Between the IOUs**
   - **Between the IOUs and the munies**
   - **Between the IOUs and merchant generators**
   - **Between the IOUs and neighboring regions**

We used to call these spot trades "economy energy" purchases and sales.

Before restructuring, we strongly urged the IOUs to take advantage of these opportunities to lower total costs. We penalized them if they failed. Whether or not you want to reform your markets, restructuring has left us an efficient mechanism for implementing these trades. Keep it. Use it.

Other actions that make sense, no matter what

4. Make the SO -- the regional dispatcher -- independent again.

- **Even in a fully re-regulated regime, a regional dispatch SO that handles bids from different parties, arranges an economic dispatch and settles economy trades will realize more gains from trade for California if:**
  - ***It functions in an impartial manner***
  - ***All parties perceive it to be impartial, and have confidence in its fairness***
  - ***It has FERC approval***

Somehow, you need to put the "I" back in the "ISO." This means:

- **You have to pay attention to the governance issue.**
- **You have to sit down with FERC and talk about how to do this.**

Other actions that make sense, no matter what

5. Re-establish the integrity of the ISO settlements.

Parties will increasingly avoid dealing with California, and avoid beneficial trades with California, if they perceive that settlements for these trades are unreliable or biased towards certain parties.

- **California consumers will lose the benefits of trade that, during the regulated years, the utilities relied on to lower costs.**
- **You can't have a single buyer, even the State of California, seen to be manipulating settlements.**

Every legitimate DWR interest can be accomplished at arms length, as a party through the ISO, without trying to control the SO.

It's time to get DWR out of the control room and out of the settlement function, and let the ISO function as it should.

Other actions that make sense, no matter what

6. If you want to make beneficial trades, get the prices right.

If you allow economy sales into California, or economy purchases from California, or if schedules have imbalances, you have to price the dispatch correctly.

- **You can't settle an imbalance that's covered by the ISO's dispatch at some average price or administratively determined price.**
- **Logically, you should settle any purchases and sales through the dispatch using marginal costs.**
  - ***Pricing at marginal costs reflects the impact of increasing or decreasing generation in the ISO dispatch to handle the imbalance.***
  - ***We know how to do marginal cost pricing of a dispatch.***
- **So settle bid-based economy purchases and sales at marginal cost**
- **And settle transaction imbalances at marginal cost**

Other actions that make sense, no matter what

7. You'll need good locational pricing, because your grid is congested. Marginal cost pricing of the dispatch will also price congestion and redispatch properly. This is called "locational marginal cost pricing."

When the ISO has to redispatch to relieve congestion, the marginal cost can differ at different locations.

- **If you allow beneficial trades, you need to recognize and use these different marginal costs at different locations.**
- **If you don't use locational marginal cost pricing, you'll either**
  - ***Subsidize the other traders (and they'll ask for more), so you have to limit trading, or***
  - ***They'll subsidize you (but they won't do that unless forced to).***
- **Once again, we know how to do this, and it works very well.**
- **Half of the United States will be using this within three years.**

Other actions that make sense, no matter what

8. Real-time pricing makes sense.

Whether you decide to reform your markets or attempt to re-regulate (whatever that means), Real-Time Pricing will provide benefits to California.

- **You will still want workable ways to get demand-side responses.**
  - ***In markets, you need the demand-side to mitigate prices***
  - ***In regulated regimes, you need demand side response to lower costs that drive rates***
  - ***Either way, you need demand-side responses to lessen shortages***
- **If you want appropriate levels of demand-side response, consumers need to face prices/rates that reflect the marginal cost of what's actually happening, as it happens. Real-time pricing does that.**
- **You just spent $35 million to get the metering in place. It's point was to allow RTP. Use it.**

Can California Rejoin the US?

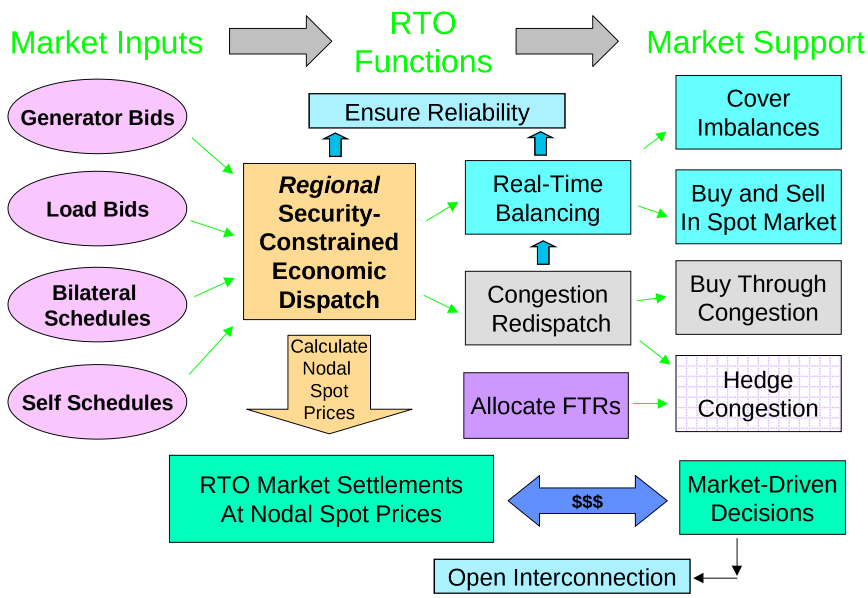
A large part of the Eastern Interconnection is moving towards a standard market design.

- **The elements are well defined**
- **We know what to do; the design has a proven track record**
- **Everything suggested here, also fits there.**

California has not decided whether to reform its markets or abandon them. However, if you look at the picture of the standard market design, almost every feature is something that has a direct analogue in a re-regulated regime. To a large degree, both choices require the same elements.

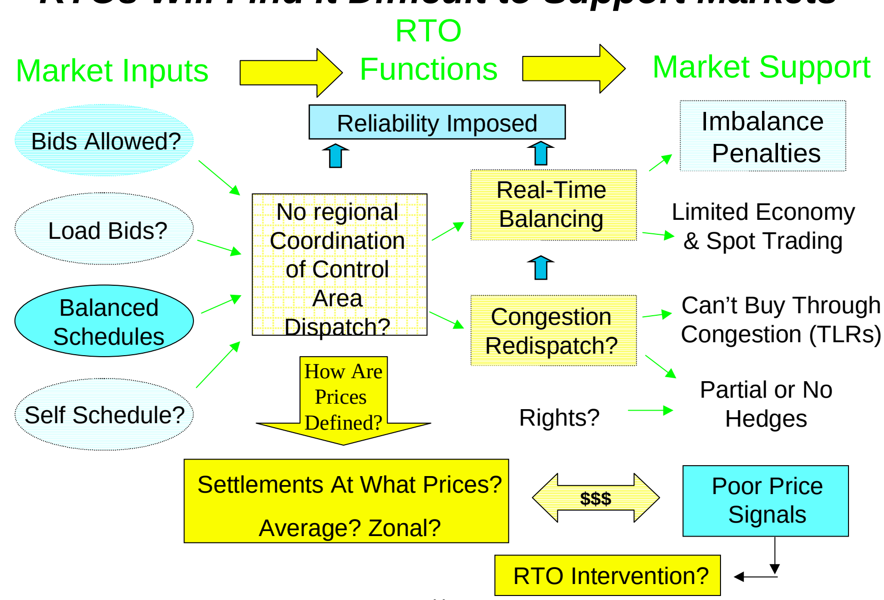
- **Everything suggested here makes sense, no matter what you decide**
- **It's all been shown to work well and to work together**
- **Doing these things will start to show benefits immediately**
- **You can then think about the choice with less risk, more confidence**

How An RTO Dispatch Supports a Regional Market While Ensuring Reliable Operations



10

Without Regional Dispatch or Congestion Pricing RTOs Will Find It Difficult to Support Markets



11

Revised: September 17, 2001

A Standard Market Design for Regional Transmission Organizations

John D. Chandley[1]

It is time for the Federal Energy Regulatory Commission (FERC) to define the principles for a standard electricity market design and to begin consistently applying those principles to the market rules now being developed by Regional Transmission Organizations (RTOs). FERC's recent RTO orders make it abundantly clear that the fundamental purpose of forming independent RTOs is not merely to "operate the grid," nor only to ensure non-discriminatory access to essential grid facilities and services. In addition to these undisputed RTO responsibilities, an essential function of RTOs is to create and operate RTO-coordinated markets.

Competing generators, customers/buyers and traders will use RTO-coordinated markets to support bilateral contracting and a wide range of commercial market transactions. The markets will also serve a broader national objective of pursuing the public interest in the benefits of market competition. RTOs will use these markets to maintain regional reliability through system balancing, congestion management and provision of operating reserves and security services, thus assuring a safe and dependable electricity system and a reliable supply of electricity. The success of RTO markets is therefore critical to system reliability, as well as national economic well being.

Market experiences in various regions have demonstrated beyond debate that poor market designs can expose affected states and participants to enormous risks. Given these risks, FERC can no longer afford to avoid the hard job of defining a workable, robust and efficient market design. And because the electricity grid functions across huge interconnected regions – as one coordinated system -- the market rules that apply across the country must be fundamentally compatible, as well as technically sound. In short, if the nation is to embrace competitive markets as the appropriate structure for the electricity industry, FERC must define and articulate the principles for a standard market design and insist that RTOs use those principles to develop their market rules. This paper describes a set of foundation principles for such a standard design.

FERC's Policy Orders

FERC has been attempting for almost a decade to restructure the electric industry to foster competitive electricity markets. After pursuing this objective under the existing industry structure through Orders 888 and 889, FERC has now concluded that it must create RTOs to achieve the goal of fostering competitive markets. In Order 2000,[2] its "Millennium Order," FERC called upon jurisdictional transmission-owning utilities to create RTOs and to transfer

---

[1] The author is a member of the LECG market design group led by Harvard's William Hogan and LECG's Scott Harvey. The paper is based on electricity market design principles developed over several years by Drs. Hogan and Harvey, other members of the design group, and other international experts in market design. The paper was commissioned by Commonwealth Edison, with the support of a number of parties interested in a clear statement of market design principles. The paper does not necessarily represent the views of Commonwealth Edison or any other party. While members of the LECG design group and others contributed helpful comments on the paper, any errors are attributable solely to the author.

[2] Regional Transmission Organizations, Order 2000, 65 Fed. Reg. 809 (Jan. 6, 2000), FERC Stats. & Regs. 31,089 (1999), order on reh'g, Order No. 2000-A, 65 Fed. Reg. 12,088 (March 8, 2000), FERC Stats. & Regs. 31,092 (2000), rehearing pending. Hereafter cited as Order 2000. All page references in the footnotes are to the Order published on the FERC web site.

control of grid operations to their respective RTOs. The Order also requires each RTO to meet certain minimum requirements to support competitive markets, but more guidance is now necessary.

FERC's Order 2000 defines the minimum requirements that every RTO must meet, while suggesting (without mandating) preferred approaches for meeting each requirement. In particular, Order 2000 requires that every RTO provide at least the following support for regional markets:

- An RTO must provide a real-time balancing market and ensure that all parties have non-discriminatory access to this market. Order 2000 notes that bid-based markets coordinated by an RTO and settled at market-clearing prices provide a proven, workable approach.[3]

- An RTO must provide market-based mechanisms to manage congestion within the region and to deal effectively with transmission loop flows, rather than rely on administrative (TLR) curtailments. FERC appears to be uncertain and less prescriptive here, but Order 2000 notes that approved, working markets have used the bid-based dispatch that provides system balancing to also provide a market-based redispatch alternative to relieve congestion.[4]

- An RTO must price congestion and imbalances efficiently, so that generators and other parties have appropriate price signals to encourage efficient short-run operations and long-run investments. Here, FERC notes its prior approvals for pricing mechanisms based on locational marginal pricing.[5]

- An RTO must offer tradable transmission rights that allow parties to hedge locational differences in energy prices resulting from congestion. These rights must support efficient regional dispatch and provide efficient incentives. The language implies a need for "financial" rights that function as price hedges rather than prerequisites for grid access that might preclude alternative grid uses or undermine an efficient dispatch.[6]

Last December, in response to Order 2000, utilities and affected parties in over a dozen regions of the United States proposed new RTOs or asked to have existing Independent System Operators (ISOs) approved as conforming RTOs. FERC has now issued initial decisions on these proposals, rejecting some proposals for various reasons and granting conditional approval to others. Importantly, on July 12, 2001, FERC issued a series of RTO decisions that, instead of leaving the number and size of RTOs to be defined through discussions among the affected parties, strongly encourage the formation of four RTOs:

[3] Order 2000, at pp. 423-425 and 633.

[4] Order 2000, at pp. 333-334, 381-382.

[5] Order 2000, at pp. 332, 382.

[6] Order 2000, at 333, 489.

> We favor the development of one RTO for the Northeast, one RTO for the Midwest, one RTO for the Southeast and one RTO for the West.[7]

FERC's desire to increase the size and limit the number of regional markets reinforces the need for a standard market design. FERC appears to recognize that its market objectives could be frustrated if there were numerous regional markets, each with a different set of parochial market rules that impose barriers to inter-regional trading. Yet the same argument applies whether there are four or 14 regional markets. The simple truth is that robust inter-regional trading and competition are unlikely to flourish if neighboring RTOs use incompatible market designs, disjointed dispatch and scheduling rules and inconsistent pricing schemes.

The broader policy objective cannot simply be a set of "seams agreements" for working around the barriers created by inconsistent regional markets, even if there are only four such regional markets. Rather the policy goal should be to facilitate seamless inter-regional trading to the degree practicable. While even fewer RTOs may be too great a technical and policy challenge for now, multiple RTOs could at least minimize seams issues by implementing a set of common market design principles that allow trading to occur across inter-regional boundaries with a minimum of administrative transaction costs and disruptions. Logically, applying a standard market design is a necessary condition, though not a totally sufficient one, for achieving these goals.

Defining a Best Practices Standard

In its July 12 Orders provisionally approving the PJM RTO proposal and rejecting the New York and New England proposals in order to encourage the formation of a single Northeast RTO, FERC indicates that it expects to see the Northeast region operating under a common set of market rules based on "best practices" from the existing ISOs:

> PJM's RTO proposal can serve as a platform for the formation of one RTO in the Northeast . . . Along these lines we encourage the three ISOs to look at the best practices in all three ISOs to develop market rules for a Northeast RTO. While we would expect that PJM will be the platform for forming a single Northeast RTO, we also would expect the RTO proposal to incorporate the best practices of the NYISO and ISO New England.[8]

Similarly, in calling for the formation of a single RTO for the Southeast region, FERC also encourages the region to develop a common set of market rules based on best practices:

> The [congestion management] mechanism ultimately proposed by Grid South to replace the interim mechanism should either implement the best practices from among the existing mechanisms currently in use by other grid operators or explain why its proposal is superior to the industry's existing best practices.[9]

[7] See, e.g., FERC, "Order Provisionally Granting RTO Status" (for the PJM RTO), issued July 12, 2001, in Docket No. RTO1-2-000, page 2. Hereafter, cited as PJM RTO Order.

[8] PJM RTO Order, page 13.

[9] FERC, "Order on Compliance Filing and Status Report," issued July 12, 2001 in Docket Nos. RTO1-74-002 and RTO1-74-003, page 23.

FERC has not determined what these "best practices" include. While Order 2000 lists several minimum requirements for an RTO's market support, it gives flexibility to each RTO in developing market rules that satisfy the Order's minimum requirements. And on approaches to congestion management, the Order declares FERC to be open to innovation and experimentation. However, there may only be a narrow range of design choices within the "existing mechanisms currently in use." The most obvious choices include:

(1) *The current pre-ISO approaches*. In regions that do not yet have independent system operators, trading must rely on utility controlled dispatches, restricted imbalance services, contract-path transmission scheduling and (in the Eastern Interconnection) administrative curtailments ("unscheduling") using Transmission Line Loading Relief (TLR) procedures. There are no spot markets, no ability to buy through congestion, and no workable systems for tradable transmission rights

(2) *The California market design*. This uses an ISO to coordinate bid-based spot/balancing and congestion management markets, while using zonal pricing to price congestion and spot market transactions. The ISO offers inter-zonal transmission rights (a form of flowgate rights) that function more or less as financial hedges against inter-zonal congestion charges.

(3) *The Northeastern market designs*. These markets also use ISOs to coordinate bid-based spot/ balancing and congestion management markets (as well as day-ahead energy and transmission markets), but they use nodal locational marginal pricing to price congestion and spot market transactions. The ISO offers point-to-point financial transmission rights that function as hedges against congestion-based spot prices and point-to-point congestion charges. Parties willing to pay the marginal cost of redispatch can buy through congestion.

(4) *A "hybrid" market design*. These might use some combination of the Northeastern designs for real-time congestion management using nodal pricing, and flowgate transmission rights to help manage or hedge congestion in forward periods. There are no designs of this type currently in operation, but variations are being considered by some emerging RTOs.

Order 2000 has already rejected the existing contract-path/TLR approach as unequal to the challenges of supporting a competitive market. It also seems unlikely that FERC would readily accept another design premised on the uniform congestion pricing schemes that it has already found to be problematic in PJM (1997) and New England (1999) or the inter-zonal/intra-zonal distinctions that it found to be fundamentally flawed in California (2000). There are no operating markets using any of the proposed "hybrid" market designs at this time, though several are being considered. While this would appear to substantially narrow the choices, FERC has been careful not to mandate any particular approach:

> . . . we reiterate that, while LMP is an acceptable approach, the Commission does not prescribe any particular congestion management method. Order 2000 grants RTOs considerable flexibility to propose a congestion pricing method that is best suited to each RTO's individual circumstances.[10]

[10] PJM RTO Order, page 19.

Learning from Experience

A reluctance to be more prescriptive may have been appropriate when FERC first approved the initial (and at that time, still untested) market designs for each of the currently operating ISOs. But over the last four years, FERC and the operating ISOs have gained valuable, sometimes costly experience in implementing different market designs. Much has been learned from market operations in PJM, New York, New England and California about what works, what doesn't, and why, as well as the shortcomings of trying to foster markets in regions without the market coordination provided by independent regional system operators. Despite continuing experiments, there may be an emerging consensus that certain design features are not only desirable but also essential, while other approaches, once thought to be attractive or simpler, have proved to be more complicated than expected, unworkable and fundamentally flawed.

Several markets have now shown what design theory predicted: simplified uniform and zonal pricing mechanisms cannot provide a workable framework for congestion management, fail to send adequate locational price signals, mask and facilitate the exercise of market power, and require increasingly more intrusive forms of intervention and restrictions to counteract the effects of inefficient price incentives.[11] At the same time, three years of planning and two years of unsuccessful pilots have shown that uncoordinated (bilateral) market redispatch cannot get the job done, because the complexities of network interactions require effective coordination through the dispatch arranged by a regional system operator.[12]

In contrast, during the same time frame, coordinated, bid-based spot markets with market-based congestion management have been operating more or less successfully in PJM and New York, applying principles of least-cost security-constrained dispatch, locational marginal pricing and tradable point-to-point financial transmission rights. Faced with the need to reform its market rules and uniform pricing scheme, New England has chosen to adopt a standard design founded on these same principles, and other regions are now considering similar approaches. Within this framework, generators have incentives to follow the dispatch required to maintain reliability, decentralized bilateral trading flourishes in ways that support reliable operations and is supported by open access to the spot market, transmission customers can elect to buy through congestion without cross subsidies, parties are compensated at market prices for providing congestion redispatch services, and investors have at least begun to see forward price signals that encourage market-driven investments at the right locations (though more might be done to improve longer-term price transparency and investment incentives).

This paper draws from these and other experiences a set of foundation principles for a "Standard Market Design." With the experience now available, the lessons learned indicate that FERC

[11] Many lessons learned from simplified zonal systems would apply as well to some of the proposals to use flowgate transmission rights, which raise issues similar to those that arise with zonal pricing and inter-zonal congestion rights. For example, the concept of defining "commercially significant flowgates" (CSFs) and offering rights only for these CSFs is analogous to the notion that important and frequent congestion can be predicted so as to define zones and inter-zonal rights, but presumably unimportant and infrequent "intra-zonal" congestion can be managed and priced with less care and no rights. Simplifying assumptions that ignore and fail to price grid realities have been the source of persistent incentive problems in several markets. These issues are discussed further in later sections of this paper.

[12] Congestion Management Working Group of the NERC Interface Committee, "Final Report on the NERC Market Redispatch Pilot," November 29, 1999, filed with FERC on December 1, 1999.

should establish the basic design principles that have worked successfully in PJM, New York and elsewhere as a Standard Design and the default "best practices" approach for other regions. That is not to say that the PJM or NY rules cannot be improved, nor does it mean that FERC should impose on every RTO the entire set of PJM or New York Market Rules, associated software and business protocols, though that shortcut is certainly available for those who choose it.[13] Rather, there is a set of core principles and essential elements that underlie these markets (and successfully operating markets elsewhere), and these provide the foundation for a Standard Market Design. With this as a proven, workable foundation, important details can then be defined and tailored to any special characteristics of each region. The Standard Market Design could serve as FERC's best practices benchmark for RTO market design and rule development.

Given the risks of poorly designed markets, arguments in favor of continued experimentation with departures from the core principles should bear a heavy burden of proof. Other innovative approaches can and should still be considered, but the proponents of an untried approach should be required to demonstrate that it is theoretically sound and likely to work, while posing minimal risks if it should fail. Equally important, and consistent with FERC's directive to meet a "best practices or better" standard, untried approaches should be imposed only if they are likely to improve on the Standard Design principles by more effectively achieving the public interest in an efficient competitive market.

The design principles offered here are neither new nor untried. The fundamentals have been described in numerous technical and policy papers over the last decade or more,[14] and the basic principles underlie the market support requirements of FERC's Order 2000. Equally important, these principles provide the foundation for successful markets in a dozen regions of the world, including workable markets in the Northeast. A successful, theoretically sound and workable model is available; it simply remains to restate its principles and concepts and apply them broadly.

The Task of Market Design

The policy justification for creating competitive electricity markets is to harness the incentives and dynamics of competition in support of efficient outcomes that enhance the public welfare. Where competition has been successful, it has led to more efficient use of resources, market-driven investments that better allocate risks and rewards, valuable innovation in products and services and more efficient prices. But restructuring a highly complex and long regulated

[13] Because of their different starting points, not everything in PJM or New York is readily transferable to other regions, and work remains to be done in PJM and New York as well, especially on the mechanisms for inter-regional scheduling and dispatch coordination, inter-regional transmission rights, associated pricing and settlement mechanisms, and a consistent and effective regional approach for assuring supply adequacy.

[14] See, e.g., F.C. Schweppe, M.C. Caramanis, R.D. Tabors, and R.E. Bohn, *Spot Pricing of Electricity*, Kluwer Academic Publisher, Norwell, MA, 1988; Larry Ruff, "Stop Wheeling and Start Dealing: Resolving the Transmission Dilemma," *Electricity Journal*, June 1994; Don Garber, William Hogan and Larry Ruff, "An Efficient Wholesale Electricity Market: The SDG&E Pooling Proposal," *Electricity Journal*, September 1994; William Hogan, "Coordination for Competition in an Electricity Market," Center for Business and Government, Harvard University, March 2, 1995; W. Hogan, "Regional Transmission Organizations: Millennium Order on Designing Market Institutions for Electricity Network Systems," May, 2000.

industry into one driven by market principles is extremely difficult, and efforts are not always successful. A poorly designed restructuring effort in an industry as complicated and essential as electricity can be disastrous, as reports from California and elsewhere attest. Such failures serve as a warning that successful electricity markets do not happen by chance or neglect, and they may be especially elusive if careful attention is not paid to the fundamentals of market design, market pricing and compatible regulatory support and oversight. Well-designed markets may well be self-correcting, but poorly designed markets and incompatible supporting institutional arrangements are clearly not. Indeed, they can become self-perpetuating.

A key premise of successful competitive markets is that they work through the interaction of private, decentralized trading and investment decisions. An effective electricity market would therefore seem to depend not only on having many buyers and sellers to ensure competitive outcomes, but also on allowing substantial commercial freedom to market buyers, sellers and various types of traders and risk managers. Trading rules would then allow the participants the freedom to fashion and implement various trading and risk management arrangements with each other, at prices to which they mutually agree, in pursuit of their respective commercial interests.

In contrast, the complexities of electricity networks require a degree of centralized coordination over system operations. Since the need for central coordination is always present, the imposition of central coordination would, at first glance, appear to be in conflict with the desire for commercial flexibility and decentralized trading mechanisms. However, this need not be the case in a well-designed market.

The interaction between the need for central coordination to achieve system reliability and the desire for decentralized commercial freedom and flexibility describes the arena for electricity market design. The task is to fashion rules for market operations and pricing that accommodate and reconcile the dictates of reliable operations and the need for commercial flexibility.

Successful electricity markets have confronted and solved this problem by explicitly recognizing the need for coordination, allowing the system operator to coordinate the short-run spot markets associated with maintaining reliability, and then consistently applying market principles and efficient pricing to the necessary coordination services and products.[15] As in any market, getting the pricing principles right is the key. The challenge in designing an electricity market is to define prices that accurately reflect what the system operator must do to maintain reliability – from calling upon more (or less) expensive generators to balance the system to redispatching units out of merit order to relieve congestion -- and thus accurately signal to grid users the true economic consequences of the grid use choices they make.

---

[15] Ruff, Larry, "Competitive Electricity Markets: One Size *Should* Fit All," *Electricity Journal*, November 1999, pp. 20-35. As Ruff notes,

> Every electricity market in which there has been reasonably effective and efficient competition among multiple sellers and multiple buyers has had a more or less independent system operator (ISO) using some sort of spot market to coordinate the physical operations of the system. Such an integrated dispatch/spot market process is the only practical way to price or internalize the complex network externalities that have traditionally justified vertically integrated monopoly utilities. The resulting spot prices and the sophisticated financial contracting they support create consistency between reliability objectives and commercial objectives.

Efficient price signals consistent with reliability then provide a stage upon which commercial freedom can play. Bilateral contracts and other decentralized trading mechanisms, self scheduling and self commitment can be freely arranged and implemented across a complex, interconnected grid without concern for cross subsidies or conflicts with reliability. Flexible dispatch/scheduling rules can accommodate both coordinated, bid-based trading and bilateral or self-nominated schedules, while flexible net accounting rules can settle transmission schedules, bilateral imbalances and spot trades at internally consistent market prices. The result is to place bilateral and spot trading on an equal footing, leaving the mix between long- and short-run options, and between decentralized trading and coordinated trading, to be decided by the market.

In contrast, unsuccessful markets have failed or performed poorly primarily because they took the opposite course, artificially separating – or attempting to separate – the necessary coordination functions from the market, limiting market access to the system operator's essential services, and then failing to price the necessary coordination functions at efficient, market-clearing prices. The predictable result has been that both markets and system coordination functions performed badly, requiring substantial intervention in commercial choices, administrative restrictions and penalties (in lieu of market prices) to enforce behavior, and re-centralized investment decision-making to preserve reliability.[16] Explaining the logic behind these observed results and designing RTO markets to avoid the same pitfalls are why a standard market design is most needed.

While faulty market designs can have extremely serious consequences for parties in the affected region, they are not easy to fix. Parties will respond logically to the perverse incentives, and as soon as advantaged parties learn to exploit the inefficiencies, achieving necessary reforms becomes difficult. Market participants advantaged by the rules have no incentive to support (and may actively oppose) market self correction, while regulators, on whom the burden of reform must now fall, are ill prepared to do the job because there is no standard design framework and no coherent set of adopted design principles from which to derive the needed reforms. Faced with pressure to "do something," regulators may impose *ad hoc* patchwork approaches that not only fail to solve the underlying problems but may make conditions worse or sow the seeds for future problems. Eventually, the need to address flawed concepts in one area requires more comprehensive and lengthy redesign in others, as both New England and California reform efforts have shown.

The common lessons derived from these experiences suggest that FERC needs a standard design framework based on proven, workable approaches and a coherent set of internally consistent market rules.[17] The remainder of this paper offers a suggested framework and then describes the foundation principles for a Standard Market Design for electricity.

[16] See, John D. Chandley, Scott M. Harvey and William W. Hogan, "Electricity Market Reform in California," Comments in FERC Docket EL00-95-000, November 22, 2000.

[17] State regulators would also benefit from a Standard Design, as well as a compatible set of guidelines for restructuring at the retail level. Incompatible efforts at the wholesale and retail ends of the same market can create huge risks for market participants, affected utilities and consumers. This paper does not address state regulatory issues except to note the obvious need for a complementary state regulatory framework.

System Operations, the Dispatch and Market Design

In electricity the laws of physics dictate certain essential features of system coordination and operations. In the United States, moreover, transmission interconnections create vast, inter-regional systems that must be coordinated and operated under common, consistent procedures and within narrow tolerances. In extreme cases, failure to continuously observe these limits can jeopardize an entire interconnection – up to half the country – in a matter of minutes. Even far less serious system failures can result in costs in the billions of dollars from damaged equipment and lost operations.[18] Some degree of centralized coordination is therefore essential to maintain reliable operations of an electricity system.

The system operator's dispatch plays a pivotal role in integrating short-run markets and reliability. To ensure reliability, each system operator uses the dispatch of generation to maintain system balances at constant frequency and acceptable voltage levels, to control inter-regional flows and to manage congestion within its area of control. These same functions also provide the basis for much of the RTO's market coordination and support. In a market-based structure, the system operator uses voluntary price offers and bids submitted by participants to arrange the dispatch that, along with regulation units, balances the system and manages congestion, while ensuring the availability of operating reserves to deal with various contingencies. The system operator then uses these price offers and bids to define market-clearing prices for the real-time balancing market and market-based congestion redispatch – the same markets that Order 2000 requires every RTO to provide.

In real time, the coordination of system operations to maintain reliability and the coordination of short-run markets for balancing and congestion management converge in an integrated dispatch process, driven by the price offers and bids submitted by market participants for that dispatch. Hence, the starting point for a workable market design is the recognition that real-time markets for balancing and congestion management are based on the system operator's bid-based dispatch.

A workable electricity market therefore requires a multi-function RTO to be both the "system operator" and the short-run "market operator." Both functions center around the real-time dispatch, driven by participants' voluntary offers and bids and the need to balance the system within its security limits. If the RTO also operates short-run *forward* markets, as the ISOs do in the Northeast, these markets also derive from participant offers and bids that are used to define a day-ahead and/or hour-ahead dispatch that balances the system and honors all grid constraints at the least as-bid cost for that forward period. The RTO can then define market-clearing prices to pay those who offer energy to the dispatch and to charge those who take energy from the dispatch. RTO-coordinated real-time and short-run forward markets arise naturally from the fact that while the ISO is using the offers/bids to arrange a dispatch to achieve reliable operations, the participants are using the bid-based dispatch to buy and sell spot/imbalance energy (and spot transmission) at market prices to support their commercial objectives.

[18] Several now familiar factors dictate the need for coordinated system operations: the speed with which electrical energy travels, the difficulties of storage, and the resulting need for virtually instantaneous balancing between electricity production and consumption; the degree of interconnection across vast geographic regions, and the need for close coordination to avoid the near instantaneous effects of failures in one region on reliability in every other interconnected region; and the absolute physical requirements for maintaining voltage support, constant frequency and stability within extremely close tolerances across the system for both safety and reliability purposes.

The RTO's market coordination functions are thus so closely related to the dispatch/redispatch functions of the system operator that it is meaningless to distinguish the "market operator" from the "system operator." The RTO must be both. As poorly-functioning markets have shown, artificially separating these functions, imposing arbitrary limits on participants' ability to use a bid-based dispatch as an open spot market, or restricting the system operator's ability to use market mechanisms to arrange an efficient (least-cost) dispatch for balancing and congestion management, create both reliability problems and market inefficiencies.[19] In well-designed systems, the independent system operator (ISO) and the independent market operator (IMO) are the same entity.[20]

System/Market Operations and Independence

If an electricity market is to function both efficiently and without discrimination, the entity that coordinates the short-run markets, controls physical access to the grid, dispatches the system and coordinates the provision of essential grid services must do so in an unbiased manner. In the United States and elsewhere, this principle is embodied in the notion of an independent system operator, whose governance and conduct must be free of undue influence from those with commercial interests in market outcomes. If the RTO were not truly independent, participants with commercial interests could exert improper influence and/or obtain preferential treatment to achieve commercial advantages for the sellers, traders or buyers they represent.[21]

The need for an unbiased or "independent" system/market operator applies no matter how the mechanisms for grid ownership are structured. An RTO formed around a single "transco" needs its system/market operator to be independent for exactly the same reasons as an RTO formed around several "gridcos," and the need does not change whether the grid ownership is structured to be "for profit" or "non-profit." No matter who owns the grid, or how the grid owners/investors recover their investments, the RTO's system/market operator must be impartial, independent from market participants, and must perform the essential coordination functions in an unbiased manner. To ensure this impartiality, the system/market operator may need to be prohibited from engaging in any commercial activities that might affect its ability or incentives to perform its functions impartially. Because market coordination and operations can affect the value of any infrastructure investments and thus materially affect the incentives for market-driven investments, those entities eligible to undertake such investment should be separated, or at least walled off, from the entity that performs the ISO/IMO functions.

Efficiency and Economic Dispatch

Each RTO must perform its essential dispatch and related market coordination functions in an efficient manner. This goal is best exemplified by the familiar principles of a security-

[19] Hogan, William, "A Wholesale Pool Spot Market Must Be Administered by the Independent System Operator: Avoiding the Separation Fallacy," *The Electricity Journal*, December 1995, pp. 26-37.

[20] The new Ontario market makes this identity explicit by calling the system operator the "independent market operator" (IMO). However, the same concept lies at the foundation of every working electricity market, no matter what the coordinating entity is called.

[21] Allowing an important buyer to gain preferential access to an RTO's system and market coordination activities is no more acceptable than allowing preferential access to selected sellers or traders. It is equivalent to allowing a vertically integrated utility to use the dispatch preferentially to advantage its preferred customers or its own generation and contracting activities.

constrained economic dispatch. Security-constrained economic dispatch is common to most systems and is well understood by system operators, whether they work for an ISO/IMO or utility-run control centers. To achieve an economic dispatch, the system operator uses each plant's operational and economic inputs to arrange a dispatch that simultaneously balances the system and meets all security constraints, and does so at the lowest economic cost, given the cost parameters available to the system operator.[22]

In a system in which generation is independently owned, and the market is coordinated by an independent system/market operator rather than a utility system operator, the RTO obtains the necessary dispatch inputs in the form of voluntary generator offers and load bids. Parties that do not wish to be subject to the system operator's dispatch can self schedule (or bid extreme prices that assure their dispatch preference), and they become price-takers. For those who choose to participate, the offers and bids indicate each party's willingness to sell or buy different quantities of energy at various prices through the dispatch during each dispatch interval. The RTO can then use these indications of market preferences to do what it must do for reliability – arrange a dispatch to balance the system, manage the flows between regions, relieve congestion and honor all security constraints – all at the lowest as-bid cost. By defining market-clearing prices for the energy offered and provided through the dispatch, the RTO can offer the participants an efficient means to buy and sell energy to support a wide range of commercial transactions – from decentralized, bilateral transactions to coordinated spot transactions -- while providing efficient price signals that are consistent with short-run operations.

Ensuring an economic dispatch is critical because the dispatch plays such a central role in supporting an efficient market. In addition to providing essential balancing and constraint management, the dispatch is the physical "provider of last resort" in all electricity markets.[23] No matter what the load is, or how many customers have selected alternative providers, the dispatch ensures that all loads are served, so it is essential that they be served at the lowest as-bid cost. In addition, the real-time spot market that flows from this dispatch provides the support for other market arrangements. For example, real-time spot prices (or day-ahead prices, if there is an RTO-coordinated day-ahead market) provide a reference for writing forward contracts. If the dispatch upon which spot prices are based is not economic, the higher costs are likely to be reflected in forward contracts as well. The spot prices also effectively eliminate the problem of liquidated damages when either contracting party fails to perform (i.e., to generate or consume) as expected. It does not matter whether the supplier injects more or less energy, or the buyer withdraws more or less energy than anticipated by the contract schedules, because the resulting

[22] FERC encourages RTOs to develop incentives to promote more efficient grid operations. The real-time dispatch is arguably an RTO's most important function. Rules that allow and encourage an RTO to arrange and implement this dispatch as efficiently as possible are necessary to achieve the goal of efficient grid operations. Indeed, it is hard to imagine how an RTO could achieve efficient grid operations without an efficient security-constrained economic dispatch. Once again, the RTO dispatch and the manner in which it is arranged and priced are fundamental to an RTO's success.

[23] The function of the "financial" provider of last resort could be met by some designated entity or entities that would be responsible for managing price risks on behalf of uncontracted ("default") customers, implying a continuing regulatory oversight of such entities. Alternatively, this function could be met – or eliminated -- by passing the RTO's spot prices through to the end-use consumers, *provided* they have effective and readily available means to hedge spot price volatility, a condition not likely to be fully satisfied when new markets open. Eventually, however, consumers who wish to avoid the uncertainty of spot prices can arrange their own risk management by choosing to enter into contracts with any retailer or entity that can provide risk management services. Under this arrangement, no official provider of last resort need be specified.

deviations or imbalances are simply met or absorbed through the dispatch as it automatically balances the system, and these imbalances can then be easily settled at the spot prices.

Open access to the RTO's imbalance/spot market also allows contracting parties to avoid the burden and expense of precise or even approximate load following. Aggregate system balancing is a reliability necessity, and that is what the dispatch provides. But since the dispatch is open to all users, mandating individually balanced schedules is not needed and may not even be practical in many cases, let alone commercially desirable. Parties are free to match their schedules if they choose, but if they can't, or choose not to, each schedule's imbalances are simply supplied or absorbed in the aggregate by the RTO's real-time dispatch, with the implicit purchases and sales by each party settled at the market-clearing spot prices that flow from that dispatch. With an open spot market, moreover, generators have a ready market for their uncontracted output, and loads and retailers have a reliable source and dependable market to obtain energy to meet their uncontracted demand or cover their contract obligations. An open spot market thus facilitates entry by new suppliers and retailers.

Restrictions on access to the dispatch-based spot market are therefore as misguided as restrictions on forward contracting; the options are complementary and both are essential pieces of a workable market. If the dispatch/spot prices are efficient, arbitrary percentage limits on spot transactions versus forward transactions are not needed and only serve to restrict commercial freedom and raise overall market costs. The key is to get the pricing right and let market choices and commercial needs define the appropriate mix between forward contracts and coordinated spot trading through the RTO's dispatch/spot market.

A Regional Spot Market Requires A Regional Dispatch

A regional spot market to support regional trading requires a regional dispatch. Most regions outside the Northeast (and California) are currently operated through smaller local control areas that do not allow open access to their dispatches nor market-priced spot markets. In these regions, important transition steps must be taken to move from local control area dispatches to a regional dispatch that is coordinated and priced by the RTO as an open spot market and effective tool for regional congestion management. Initially, this may require a hierarchical approach in which the RTO implements its regionally coordinated dispatch through existing local control areas.

New RTOs must eventually go through this dispatch integration process, because the reach of a coordinated dispatch largely defines the scope for regional balancing, congestion and loop-flow management, consistent pricing and transmission rights. The broader the scope of this dispatch, the broader the reach of the regional market.

With a regionally coordinated dispatch as the foundation, an RTO can define related market rules through consistent application of economic and market principles and careful observance of the requirements for reliability. Bid-based, RTO-coordinated markets can be designed to be open to all participants on a non-discriminatory basis. The RTO can define market-clearing prices to settle all purchases and sales in its coordinated markets, leaving participants free to pursue a wide range of commercial transactions and choices, with price signals consistent with the RTO's reliability mandates. These features define the essential components for RTO market support, as required by Order 2000, and provide the foundation principles for a best practices Standard Market Design.

Foundational Principles for a Standard Market Design

- Each RTO must have a system operator that coordinates system operations through a regional, security-constrained, economic dispatch. The RTO uses this regional dispatch to maintain system balances and adjusts the dispatch to relieve congestion and keep flows within all security limits, at the lowest as-bid cost. RTO coordination of the region's dispatch is therefore the cornerstone for RTO market support.

- An RTO would use market mechanisms to support its essential system operation functions, first by accepting price offers and bids from those generators and loads eligible to participate in the RTO's security-constrained economic dispatch, and second by defining market-clearing prices for the energy bought and sold through the dispatch. The RTO would thus operate an open, bid-based spot market.

- As its market coordination capabilities mature, an RTO would use similar market processes to coordinate the acquisition and deployment of resources for regulation and operating reserves. The RTO would coordinate these markets with its real-time energy market to achieve consistent, market-clearing prices across all of its coordinated short-run markets. RTO-coordinated forward markets (e.g., day-ahead) could be added, if desired by participants, to provide further price certainty for energy and transmission prices and a means to deal with startup/commitment risks.

- Market participants – generators, loads and traders – would participate in the provision and pricing of these services through the voluntary submission of quantity/price offers and bids to the RTO, for its use in system balancing and spot trading, constraint/congestion management, and maintaining adequate regulation and reserves. Under a consistent pricing and settlement framework to prevent cost-shifting, "self-provision" of these latter services would also be accommodated.

- The RTO would function as both an independent system operator and an independent market operator. RTO operational and market coordination functions that center around the dispatch are inseparable in real time.

- To ensure unbiased (non-discriminatory) and efficient operations, these integrated functions would be performed by an entity with no commercial interest in market outcomes. Operating an open, efficient market in an unbiased fashion would be regarded and protected as an activity with public interest objectives. These involve operating fair, open, transparent and efficient competitive markets for purposes of promoting economic welfare and an efficient allocation of resources, while ensuring reliable system operations. RTO activities, rules, procedures and affiliations inconsistent with these public interest objectives would not be permitted.

These principles provide the initial foundation for the market support functions each RTO must provide. In addition, FERC's principles of open access, driven by the twin policy objectives of economic efficiency and non-discrimination, require additional foundation principles for a Standard Market Design:

- Market participants would have open, non-discriminatory access to the essential products, services and markets coordinated by the RTO. For example, there would be no arbitrary limits on the ability of parties to participate in the RTO's bid-based dispatch. There would be no limits on their right to use the resulting balancing and congestion management markets that flow from the dispatch coordinated by the RTO.[24] Open access to balancing markets means an open spot market for buying and selling energy at market-clearing prices.

- Whatever services the RTO must provide, it should provide efficiently. For example, there should be no restrictions on the RTO's ability to arrange an economic dispatch – i.e., at the lowest-as-bid cost -- given the price offers and bids submitted by the participants and the security constraints that the RTO must honor in the dispatch.

- Whatever services the RTO must provide, it should price efficiently, using (wherever practical) market-clearing prices derived from market bids and principles of marginal cost pricing.

- Consistent with efficient pricing, the RTO would price the dispatch (and congestion redispatch) as accurately as possible, so that the resulting price signals reflect the effects of congestion and are consistent with what the ISO requires parties to do to relieve congestion and maintain reliability.

These last principles require that market prices be based on principles of marginal costs and be consistent with security-constrained economic dispatch. Prices consistent with the dispatch would therefore reflect the marginal cost of any redispatch required to relieve congestion and accommodate each transaction. These are the principles that underlie nodal spot pricing, or locational marginal pricing.[25] An RTO would apply the principles of locational marginal pricing to define market prices for spot energy and transmission.

- The price for imbalance/spot energy at each location on the grid – the "nodal price" -- would be the locational marginal price, which is defined as the incremental cost in the dispatch of meeting an increment of load at each location, given the price offers and bids,[26] the actual dispatch, and the grid constraints affecting that dispatch.

    - Generators providing energy for the dispatch would be credited in the RTO settlements for their injections at the nodal prices for their respective locations. Additional injections net of bilateral schedules would therefore be sales to the spot market and would be paid the nodal prices.

    - Loads served by the dispatch would be debited in the RTO settlements for their withdrawals at the nodal prices for their respective locations. Additional

[24] Order 2000 recognizes that open access to the real-time balancing market is required to ensure non-discriminatory access to transmission. See, Order 2000 at p. 425.

[25] Nodal prices would differ by location because of the effects of congestion and losses. An initial pricing system might not reflect marginal losses (as in PJM), but complete nodal systems would (as in New York).

[26] In calculating the market-clearing prices from the dispatch, the RTO would use the offers and bids of parties that were actually following the RTO's dispatch instructions within some defined tolerance. All other parties would be viewed as setting their own schedules and would be settled as price-takers.

withdrawals net of bilateral schedules would therefore be purchases from the spot market and would be charged at the nodal prices.[27]

- Schedules for bilateral contracts would be fully accommodated by the RTO settlements. Imbalances (deviations) from bilateral schedules would be settled at the nodal prices, such that

  (1) generators injecting less (more) than their schedules would be charged (paid) the nodal price at their location for their deviations and

  (2) loads withdrawing more (less) than their schedules would be charged (paid) the nodal price at their location for their deviations.

- The price for transmission usage for any transmission schedule (bilateral transaction) would be the marginal cost of the redispatch necessary to accommodate that schedule or transaction. The marginal redispatch cost for each megawatt of a schedule or transaction is equal to the nodal price at the point of delivery minus the nodal price at the point of receipt.[28] The usage charges that apply to transactions that are "scheduled" with the RTO can be applied as well to transactions with "loop flows" across the RTO-controlled grid from transactions scheduled outside the RTO.[29]

Efficient Pricing and Market Flexibility

Order 2000 requires that an RTO's market mechanisms provide all grid users with efficient price signals that reflect congestion and expansion costs and give participants efficient incentives regarding the consequences of their transmission usage.[30] Providing efficient signals regarding short-run usage and long-run investments is the key to giving participants commercial flexibility in choosing and implementing transactions. Where these pricing rules are followed, generators and other parties can be given substantial commercial freedom to respond to the resulting price signals without significant concern that their commercial decisions will undermine reliability. Where these rules are not followed, the system operator must inevitably intervene in the market through administrative curtailments, restrictions and non-market penalties.

---

[27] In most regions, small customers do not have interval meters, so applying hourly (or shorter) nodal prices requires some averaging and the use of load profiles. Also, such loads may not yet be accurately mapped to individual buses (nodes). In these cases, loads (or Load-Serving Entities purchasing on their behalf) can be charged an aggregate nodal price, defined as the load-weighted average of the nodal prices in the area in which the loads are located. For state retail purposes, the aggregations can cover a service area or portions thereof for some transition period.

[28] For simplicity, this difference is also referred to as the "nodal price difference" or "locational price difference." The difference may be positive or negative.

[29] An RTO that uses this market-based method to price the redispatch costs imposed by such loopflows avoids the dilemma of having either to absorb uncompensated redispatch costs or to impose TLR curtailments as its only alternatives. Instead, the RTO provides a market-priced redispatch service and allocates grid use efficiently to those willing to pay the marginal costs of that redispatch.

[30] Order 2000 at pp. 382, 489.

Ample experience in markets from New England to California demonstrates the importance of applying principles of marginal cost and aligning real-time market prices with the security-constrained economic dispatch used to maintain reliability. If imbalance/spot market settlement prices are not allowed to reflect actual market-clearing prices derived from marginal costs, parties have incentives to use the imbalance/spot energy markets in ways that shift costs to other parties. Similarly, if transmission usage charges do not reflect the full opportunity costs of any usage (that is, the marginal costs of any redispatch required to accommodate a transmission schedule), parties have incentives to over-schedule transmission usage. Accommodating these schedules through redispatch would then shift redispatch costs to others and encourage more parties to over schedule, providing an incentive to increase congestion.

Where an RTO's prices do not efficiently allocate access to the grid, the RTO must use administrative restrictions to limit access to keep flows within security limits. That is why control areas that provide no mechanism to price the marginal cost of redispatch must resort to physical curtailments under TLR. For similar reasons, an RTO using inefficient pricing would also have to limit access to its imbalance/spot energy market.[31] Balanced schedule requirements and imbalance penalties (several RTOs), penalties to discourage "under-scheduling" (California), denial of "congestion buy-through" (everywhere but PJM and New York), restricted access to avoid curtailments (throughout the West except California) and curtailments under TLR (throughout the Eastern Interconnection, except PJM and NY) are all real-world examples of administrative restrictions that have been (and continue to be) used to offset the incentives of inefficient pricing rules.

Failure to use market-clearing prices that reflect marginal costs at each location also requires an RTO to use non-market side payments to persuade plants required for congestion redispatch to follow dispatch instructions. Side payments to constrained-on and constrained-off generators are needed because uniform or zonal settlement prices, which are not based on locational marginal costs, are inconsistent with participant bids and the requirements for redispatch. Constrained-off generators must be paid not to run, or else administrative controls must be placed on plant operations to prevent them from operating in response to the inefficient price signals provided by a uniform or zonal price. However, experiences in several markets show that these side payments encourage strategic bidding to maximize the side payments.[32]

[31] Experience with non-locational pricing in PJM during 1997 showed that where pricing for grid usage is inefficient, restrictions on access to the grid must be coupled with restrictions on access to the balancing/spot market. The reason is that selling energy at location A and buying energy at location B through the imbalance/spot market is equivalent to scheduling transmission from A to B. If parties are limited through one mechanism, they will exploit any equivalent mechanism in response to the same incentives. In PJM, parties implementing internal transactions during 1997 were allowed access to the spot market to support their transactions, while parties implementing exports and imports were denied access to the spot market, to avoid having the redispatch costs imposed on the PJM members. The PJM rules did not allow open access to the spot market for all parties until the ISO implemented nodal pricing and was able to settle each transaction using the nodal price that applied to its locations.

[32] Such side payments have been used, and subsequently abused, in the United Kingdom, PJM (1997), California (today) and ISO-NE – that is, in markets that do not use nodal locational pricing. The constrained-on and constrained-off bidding games first arose in the UK and were later observed in extreme forms in the California market, which now sets a floor on bids to reduce gaming incentives. PJM eliminated the problem in 1998 by moving to nodal pricing, which does not require such side payments. New England avoided some of these problems by not compensating constrained-off plants, while limiting self scheduling. New England now plans to implement nodal pricing.

In some cases, inefficient pricing rules may also encourage parties to schedule during forward periods in ways that create artificial congestion. The RTO must then pay the same parties to relieve this congestion in real time.[33]

Administrative restrictions on investment decisions are also required, because the prices fail to provide efficient investment signals. If real-time prices are inefficient, then expectations of forward prices will not provide accurate signals for future investments. For example, if real-time and forward market price signals do not clearly indicate whether new generation at a location will help solve congestion or make it worse, the RTO must intervene in new investment and siting decisions, imposing interconnection restrictions and upgrade costs on new entrants, and so on. FERC has recognized that these restrictions discourage new entry and thus hinder the development of more competitive supplies. In both California and New England, FERC rejected such restrictions and directed the ISOs to reform their congestion management and pricing rules.[34]

The universal lesson from many markets is that inefficient pricing schemes create incentives for short-run strategic behavior while encouraging participant decisions that are inconsistent with reliability, thus forcing intervention by the system operator. This necessary intervention directly limits market freedom and flexibility. Because no system operator can allow market actions to jeopardize system reliability, the system operator must eventually intervene in the market and impose administrative controls and non-market penalties to maintain reliability, discourage gaming and limit cost shifting.

---

[33] This response to poor pricing signals arises in California and is partly a function of using unrealistic "commercial models" to manage only inter-zonal congestion in the forward markets, while the ISO must use more realistic operational models to manage all of the congestion in real time. The same concept of managing only "commercial" congestion in the forward market and dealing with "operational" congestion in real time can be seen in some of the RTO "hybrid" market design proposals. The concept has been a source of persistent gaming and operational problems in California and should be avoided. See, e.g., FERC's "Order Accepting for Filing in Part and Rejecting in Part Proposed Tariff Amendment and Directing Reevaluation of Approach to Addressing Intra-zonal Congestion," January 7, 2000. In that order, FERC noted that the California scheme for zonal pricing and intra-zonal congestion management was "fundamentally flawed" and required "a comprehensive replacement congestion management approach."

[34] For California, see FERC's "Order Rejecting Proposed Tariff Revisions," in Docket ER99-3339-000, issued September 15, 1999. For New England, see FERC's "Order Conditionally Accepting Compliance Filing, as Modified, and Accepting, in Part, and Rejecting, in Part, Proposed Tariff Changes, as Modified," in Docket No. ER98-3853-000, issued October 29, 1998; also see, "Order Extending Congestion Cost Allocation Methodology on ISO-NE's Proposal to Replace New England's Uniform Pricing on an Interim Basis and Deferring other Issues," in Docket Nos. EL00-59-000 and ER00-2005-000, issued June 1, 2000. In the latter Order, FERC approved ISO-NE's proposal to retain a scheme that required that average redispatch costs be charged to parties through an uplift until it could be replaced with a system based on locational marginal pricing. In the interim, however, FERC noted at pages 6-7:

> Although we will retain the current allocation methodology for now, we will not permit socialization of congestion costs to continue indefinitely. Given the increase in congestion in New England and the significant planned generation addition, it is essential that the ISO implement a new CMS [congestion management system] that relies on market mechanisms to establish price signals that will serve to allocate constrained transmission to the highest valued users and give generation an incentive to locate in appropriate areas.

On the other hand, where energy and transmission prices accurately reflect the marginal costs of congestion redispatch, the settlement prices encourage generators to follow this dispatch. By eliminating the need for side payment schemes, nodal pricing removes the opportunity to gain commercially from strategic bidding to maximize the payments. The resulting market-clearing prices tend to reinforce reliability, rather than undermine it. Short-run market prices then provide an appropriate foundation for forward price signals for siting new generation (and loads) at locations that will help relieve congestion, rather than exacerbate it, while providing an incentive for market-driven transmission upgrades.

Efficient pricing is therefore another cornerstone in the Standard Market Design. Achieving commercial flexibility and liberating decentralized decision-making requires that an RTO price its essential coordination functions efficiently, using principles of marginal costs, and in a manner consistent with the dispatch it implements to maintain reliability. With an efficient pricing foundation that supports reliability, the Standard Design can then accommodate a broad range of commercial flexibility and market trading options. For example:

- Market rules can offer market participants maximum commercial freedom with respect to real-time operations, forward trading, and long-run investment decisions. The RTO can support such commercial flexibility without significant concerns for reliability, unfair cost-shifting or inappropriate investment decisions that have either short-run or long-run adverse implications for system security. Within a nodal pricing framework:

    - Parties may choose to participate in the RTO's coordinated regional dispatch. Generators may submit voluntary quantity/price offers to the RTO dispatch, and loads may submit voluntary quantity/price bids to the RTO dispatch.

    - Any participant may buy and sell spot energy through the RTO real-time dispatch/imbalance market and be settled at the nodal spot prices without concern that such transactions will impose unfair cost shifts or subsidies on others. Spot purchases and sales defined by marginal costs do not "lean" on the RTO or anyone else and therefore need not be discouraged, limited or penalized.

    - Generators and Load-Serving Entities may implement bilateral arrangements through either scheduling or settlements.[35]

    - Parties may submit balanced schedules or not, depending on their commercial needs.[36]

---

[35] The distinction here means that bilateral parties do not have to "schedule" their bilaterals in advance, though they may choose that option. Alternatively, they can self-schedule individually, or participate in the dispatch individually, and then inform the RTO of their bilateral arrangement only for settlement purposes, after the dispatch is completed. In the latter case, the RTO settlement nets the supplier's credits (its injections times the nodal price at its location) and the buyer's debits (its withdrawals times the nodal price at its location) to produce a net bill that reflects imbalances and any congestion-related usage charge.

[36] The freedom for both generators and loads to submit offers and bids to the RTO for its use in system balancing and congestion redispatch without regard to balanced schedules adds to the depth and competitiveness of the real-time balancing and congestion management markets. The more competitive these markets are, the more effective they are in managing the system through market prices and voluntary market choices, rather than RTO administrative decisions. Allowing loads to bid without balanced

- Parties may offer non-contracted energy to the RTO for dispatch, purchase spot energy to support a bilateral, and use the balancing/spot market to purchase energy to meet any uncontracted demand. Such purchases and sales are settled at the market-clearing spot prices.

- Any transmission customer (i.e., those with bilateral schedules) may submit transmission schedules that the RTO will accommodate through bid-based economic redispatch arranged by the RTO, as long as redispatch bids are available and the customer agrees to pay the marginal cost of redispatch to accommodate the transaction. A party may thus choose whether to "buy through congestion," based on its willingness to pay the transmission usage charge, or choose to self-curtail if it expects the transmission usage charge (equal to the nodal price difference) to exceed its willingness to pay.[37] Parties otherwise subject to TLR curtailments due to loopflow through the RTO system have the same options.

- A transmission customer may implement a schedule that provides "counter-flow," which lowers the marginal cost of redispatch, and be compensated accordingly for the value of that counter-flow.[38] Parties thus have another means to provide congestion management and be paid its market value.

- Any Load-Serving Entity (LSE) may self-schedule its own generation to serve its own loads and/or use the RTO balancing/spot market in any mix it chooses. The LSE is subject to the same imbalance/spot pricing and transmission usage charges as any other party. There is no bias in favor of or against "native loads."

- Any generator, trader, customer or LSE may determine the mix of contract and spot purchases or sales appropriate for its commercial circumstances.

- The RTO can create financial trading "hubs" based on any aggregation of nodes (grid locations where spot prices are defined by the RTO using nodal pricing), with the hub prices defined as the fixed-weighted average of the prices at the nodes included in each hub. Market participants may then freely trade at, to and from these hubs and have their energy trades and transmission rights settled by the RTO using the hub prices. Similarly, retailers can have any collection of nodal prices aggregated to facilitate settlements with loads at multiple locations.

---

schedules also provides an effective means for defining market-clearing prices during shortage conditions and mitigating the potential for market power.

[37] Within the nodal pricing framework, the opportunity cost of any transmission usage equals the marginal cost of redispatching the system to accommodate that usage, which is in turn equal to the difference in nodal prices. The RTO would continuously post the nodal prices in near real time to allow parties to exercise these options more effectively.

[38] The value of the counterflow is defined by the savings in opportunity costs or marginal redispatch costs, as defined by the nodal price difference that applies to the points of receipt and delivery for the counter-flow schedule.

- Generator investment and siting decisions can be market-driven without significant RTO direction or interference.

- New generator interconnection procedures can focus on procedures to ensure safe and reliable interconnection and not become mechanisms to restrict new entry for fear of impacts on grid congestion. Nodal prices and related congestion (transmission usage) charges, and expected forward prices based on these real-time prices, will tend to discourage siting at locations that worsen congestion and tend to encourage siting at locations that relieve congestion. Interconnection rules therefore need not discourage new entry by imposing system ("deep") upgrade costs as a condition for new interconnections.

- Through the award of transmission rights to those who fund upgrades and increase system capacity, at least some transmission investments can be market-driven, in the absence of market failure, and should be accommodated by the RTO.

Transmission Rights in a Market with Nodal Pricing

Under the Standard Market Design, the risks associated with uncertain physical delivery due to congestion are virtually eliminated by the RTO's dispatch. Whenever proposed schedules would create congestion, the RTO would provide redispatch for parties willing to pay the marginal costs of any redispatch needed to accommodate those transmission schedules, rather than ration grid access through physical curtailments based on administrative priorities. By eliminating the risks of physical delivery, the availability of an RTO redispatch option allows market participants to focus on the financial risks associated with their transactions.

In the Standard Market Design, the financial risks associated with congestion become transparent through the locational marginal (nodal) prices and associated usage charges, but parties need an appropriate financial instrument to hedge these risks. The RTO will use nodal prices to settle imbalance/spot energy purchases and sales; it will use nodal price differences between points of receipt and delivery to define transmission usage charges that reflect the marginal cost of redispatching the system to accommodate each transaction. However, parties will not know the effect of congestion on spot energy prices and transmission usage charges until after they complete their transactions and after the RTO implements the dispatch and calculates the nodal prices at each location. Spot pricing is *ex post*.

Participants will therefore need a system of transmission rights to manage the risks of price uncertainty arising from pricing congestion. This purpose can be expressed in different ways:

- Transmission rights in a nodal pricing system provide a mechanism by which a party can lock in the price of transmission – the usage charge -- in advance.

- Transmission rights in a nodal pricing system allow a party to get access across the grid to the price at another location, even though there may be congestion between that location and the party's location. [39]

[39] In a nodal pricing system, a "location" for settlement purposes can be defined as a specific bus or node or as the weighted average of many nodes. For a Load-Serving Entity, for example, its location can be defined by all the nodes serving that load, and this aggregated settlement price can be defined as the load-

- Transmission rights in a nodal pricing system provide a means by which a party can offset the impact of congestion on prices at any location.

An effective transmission right (or set of rights) for a transaction from A to B would allow a party implementing that transaction to be credited for the same nodal price difference that defines the usage charge for that transaction, so that the credit received for a right matching a transaction would exactly offset or hedge the usage charge for that transaction. A party holding that transmission right would be perfectly hedged against any A-to-B usage charge (the congestion-related differences in nodal prices) and would thus have effective access to the price at location A, the generator/source location, even during congested hours, as the following illustration shows.

Transmission Rights as Financial Hedges

Where the nodal price at generator/receipt location	= Nodal Price at A
And the nodal price at load/delivery location	= Nodal Price at B,
And the party holds a right from A to B	

Then,

The transmission usage charge	= Nodal Price at B minus Nodal Price at A
The transmission right credit	= Nodal Price at B minus Nodal Price at A

The transmission right serves to offset or "hedge" the usage (congestion) charge.

(Note that the hedge in this illustration functions whether the Nodal Price at A is higher or lower than the Nodal Price at B, because the usage charge will also change in an exactly corresponding fashion. See discussion of "obligations.")

In this illustration, the transmission rights are directional and defined from point to point (location A to location B). The RTO defines the settlement credit for the financial transmission rights using the same nodal prices that it uses to define the transmission usage charges.[40] This symmetry makes the operation of the rights transparent and intuitive, while simplifying the settlement system.[41]

---

weighted average of the nodal prices that make up that aggregate. Any trading hub can be defined in a similar manner.

[40] Revenues to fund these hedges would come from the settlement surplus collected by the RTO in its coordinated markets. The surplus arises naturally in a system that prices congestion and derives here from usage charges (differences in nodal prices) paid by transmission customers and from the fact that the aggregate revenue from spot energy prices paid by loads is greater than the payments made to generators.

[41] In theory, an equivalent credit could be defined consistent with the same nodal prices by pricing each of the binding constraints between the source and sink locations. The credit for a matching set of such financial "flowgate" rights for the exactly equivalent set of constraints would *in theory* provide the same hedging value in a nodal pricing system as the point-to-point right assumed in the illustration. However, equivalency under all conditions depends on assumptions that the binding constraints are predictable and stable, and that the flows across them are predictable and stable. If these assumptions do not hold, then the equivalency disappears. Where RTOs are considering the use of flowgate rights, special care must be taken to ensure that settlement rules do not assume equivalent credit under conditions that do not support

In addition, the settlement value of a financial transmission right between any other set of locations – e.g. a right between locations C and D -- would be defined in the same manner, as the difference in the nodal prices at those two locations. Thus, any financial transmission right (or set of rights) could be used to hedge (or partially hedge) any schedule's usage charge, and there would be no requirement that the rights match the schedules. Parties could therefore select, acquire and trade rights for risk management purposes independently from the implementation of, or changes in, their actual schedules or the RTO's economic dispatch (or congestion redispatch), and the rights would still serve the risk management function.

Order 2000 Requirements for Transmission Rights

FERC's Order 2000 requires that each RTO provide a system of tradable transmission rights that can hedge locational price differences resulting from congestion. Further, FERC requires that the transmission rights system must promote an efficient dispatch.[42] This implies that the system of transmission rights should not encourage or lock the rights holders or the RTO into an inefficient dispatch.

Nodal pricing is premised on an efficient dispatch. The RTO uses voluntary price offers and bids to arrange a security-constrained, economic dispatch. This dispatch balances the system while honoring all constraints at the lowest-as bid costs. This is consistent with the FERC directive that generators dispatched in the presence of congestion be those that can serve loads at the lowest cost.[43] Hence, the system of transmission rights for a Standard Market Design must be compatible with, and not undermine, the RTO's security-constrained economic dispatch.

Order 2000 also requires that the RTO's prices provide efficient signals regarding short-run operations and long-run investments. Under Order 2000:

> [E]very RTO must establish a system of congestion management that establishes clear rights to transmission facilities and provides market participants with price signals that reflect congestion and expansion costs. (p. 489)
>
> "[W]e will require the RTO to implement a market mechanism that provides all transmission customers with efficient price signals regarding the consequences of their transmission use decisions. We are convinced that efficient congestion management requires that transmission customers be made aware of the cost consequences of their actions in an accurate and timely manner, and we believe that this is best accomplished through such a market mechanism." (p. 382)

FERC therefore requires a system of transmission rights to do more than just hedge congestion charges. The transmission rights must also support an efficient dispatch and support a pricing system that provides efficient price signals that accurately reflect congestion and expansion costs. The combined system must efficiently allocate grid use to those who value it the most.

---

equivalency, to avoid creating inefficient arbitrage opportunities and undermining the efficient incentive properties of the nodal prices. Principles that should guide an RTO on this question are discussed below.

[42] Order 2000, at p. 333.

[43] Order 2000, at pp. 332-333.

Financial Transmission Rights in Practice

PJM and New York each provide a system of point-to-point financial transmission rights, or FTRs,[44] that function like those in the illustration above. These FTRs entitle the holder to a settlement credit equal to the difference in nodal prices (or the congestion component, where the nodal prices also include marginal losses), thus providing a means to hedge congestion-based usage charges that are defined by the same nodal price differences. Moreover, because these FTRs are financial, they do not interfere with or undermine the ISO's least-cost dispatch. This financial aspect can be expressed in different ways:

- The rights entitle the holder to a monetary credit in the RTO settlements; they do not guarantee physical access to the grid.

- Parties are not required to acquire or hold FTRs as a condition for scheduling or gaining grid access.

- Parties are not required to match their FTRs to their schedule, or to match their schedules to their FTRs.

- Parties are free to participate or not in the ISO's security-constrained economic dispatch, whether or not they hold FTRs, and the dispatch does not depend on the FTRs they hold.

- In the RTO settlements, parties receive the market value of the FTRs they hold (defined by the nodal price differences), whether or not they implement a matching schedule.

These financial attributes of FTRs help ensure that parties receive the risk management value of the rights they hold while remaining free to adjust the points of delivery and/or receipt for their schedules or participate in an efficient dispatch irrespective of the rights they hold. Parties are not discouraged from making these adjustments and have no incentive to undermine an efficient dispatch, whether or not they hold rights, and the RTO is free to arrange an efficient dispatch or redispatch without having to worry about whether the resulting flows match each party's rights. This would not be the case, however, if the rights were "physical."

If the rights were "physical," parties would need to acquire rights matching their expected transactions as a condition for gaining grid access. If they changed the points of receipt or delivery for their expected transactions, they would have to exchange the rights they held for others that matched their revised schedules. If they were unsuccessful, they would be excluded from the grid and/or risk losing the value of any rights that no longer matched their expected transactions.

[44] This paper uses the generic term FTR to mean "financial transmission right." FTRs are called "fixed transmission rights" in PJM and "transmission congestion contracts" in New York. While they are allocated differently, in all fundamental respects, PJM's FTRs and New York's TCCs are identical. An identical approach is under development for ISO-New England, where the rights are called "financial congestion hedges." A Northeast RTO would probably use the generic term "financial transmission right."

Parties that held physical rights could also exclude others from using the grid merely by holding these rights until scheduling deadlines were passed. "Use-it-or-lose-it rules" have been proposed to prevent such hoarding in proposed physical rights regimes. However, if such rules were strictly enforced, they would limit trading flexibility, potentially forcing parties to trade rights they might need later before they were certain about this need. If such rights were traded on a recallable basis to avoid losing them, and then recalled by the original party, the counter party would be left with an unreliable right and little time to hedge the risks created for its schedules.

Given these risks, parties holding physical rights would be discouraged from changing their expected points of receipt and delivery, even if it meant a more profitable transaction or a more efficient dispatch. Such changes would require parties to engage in additional trades to realign their rights portfolio with their revised transactions. At any given moment, the distribution of rights would essentially define the dispatch, making it difficult for the RTO to arrange an efficient dispatch that could require generators and rights holders to reallocate the necessary rights to accommodate the more efficient dispatch. Furthermore, an efficient redispatch to relieve unexpected congestion would be almost impossible to achieve in real time, unless the RTO simply ignored the rights.

Physical rights systems applied to an electricity network thus interfere with an efficient dispatch and efficient congestion management coordinated by an RTO. They tend to discourage more efficient operations by the RTO, while imposing additional trading costs and risks on the market that are difficult to manage.

In contrast, financial transmission rights avoid these dilemmas because they are financial instruments that can be acquired and traded independently from the physical dispatch of energy. Because financial rights do not control physical operations, retaining their value need not compromise efficient physical outcomes nor complicate the physical redispatch required to ensure reliability when the grid is constrained.

Consistency Between Financial Rights and Nodal Prices

An attractive and essential feature of the point-to-point FTRs offered by PJM and New York is that each ISO defines an FTR's settlement value using the same nodal prices used to define congestion usage charges. As shown in the illustration above, the settlement prices for FTRs and the settlement charges for congestion are based on the same nodal prices. This means that congestion charges and transmission rights that serve as congestion hedges are calculated in a consistent manner, using the same grid conditions and assumptions, and both are consistent with the dispatch for the market in which they are settled.

This internal consistency ensures that the FTR settlements do not undermine or distort any of the efficiency properties of the nodal prices and related usage charges. Parties receive efficient price signals from the nodal prices for both short-run operations and long-run investments, and nothing in the FTR settlement rules undermines the efficient signals that each party receives regarding its use of the grid.

FTRs defined in this manner fulfill all of the objectives of Order 2000. They are consistent with FERC's goals of an efficient dispatch, efficient price signals and efficient allocation of grid use. An RTO implementing the Standard Market Design should therefore offer financial transmission rights that meet these same principles.

- In conjunction with a nodal pricing system, an RTO should offer and provide settlement support for a system of point-to-point financial transmission rights.

- The FTRs are “financial,” in that they entitle the holder to a settlement credit that can offset the congestion-related components of nodal prices or the congestion-based usage charges for a corresponding transaction. The financial characteristics of FTRs are required to facilitate independent rights trading while ensuring that they do not undermine an efficient dispatch or distort efficient price signals:

    - Parties are not required to acquire or hold FTRs as a condition for scheduling or gaining grid access.

    - Parties are not required to match their FTRs to their schedules.

    - Parties are free to participate or not in the ISO’s security-constrained economic dispatch, or to schedule transmission, whether or not they hold FTRs.

- In the RTO settlements, parties receive the market value of the FTRs they hold, whether or not they implement a matching schedule.

- Settlement rules for FTRs should not undermine or distort the efficient price signals provided by the nodal pricing system. FTR settlements should use the same nodal prices that are used to define spot market settlements and usage charges, so that congestion charges and the value of congestion hedges are defined by the same grid conditions and assumptions, and both are consistent with the actual dispatch.[45]

Point-to-Point and Constraint/Flowgate Rights

The FTRs described in the previous sections and offered in PJM and New York are defined from point to point, with their values defined by the nodal prices at each point. The Standard Market Design includes point-to-point FTRs.

Some parties have indicated a desire to manage congestion pricing risks through transmission rights that are defined on a constraint-by-constraint basis. Constraint-specific rights, or “flowgate” rights (FGRs), as they are often called, could also be defined as financial rights.[46] Whether they could meet each of the foundation principles in this paper, however, remains to be demonstrated. For example, to be compatible with the Standard Market Design, and avoid undermining the efficient price signals achieved by the nodal pricing system, the flowgate rights would need to meet the following requirements:

[45] If the RTO operates a day-ahead market, the FTRs would be settled in the day-ahead market using day-ahead nodal prices. Day-ahead FTR credits would therefore be consistent with day-ahead nodal prices and usage charges.

[46] A point-to-point FTR can be viewed as a complete set of flowgate rights for all of the possible constraints that could be binding between the two grid points defined by the FTR. In theory, therefore, there is a set of flowgate rights that corresponds to each point-to-point FTR. However, this correspondence does not hold as grid conditions change, thus complicating efforts by an RTO to use FTRs and flowgate rights interchangeably. However, nothing in the Standard Design’s inclusion of FTRs would prevent private parties from acquiring FTRs and making their own markets for trading in flowgate rights.

- *FGRs* would be financial rights, and be subject to the same principles regarding their financial attributes that apply to point-to-point FTRs.

- In the RTO settlements, parties would receive the market value of the *FGR*s they hold, whether or not they implemented a matching schedule.

- Settlement rules for *FGR*s should not undermine or distort the efficient price signals provided by the nodal pricing system. To avoid undermining the efficient incentive properties of the nodal prices, settlement rules for *FGR*s should be consistent with those prices, with the value of *FGRs* determined by the same dispatch, grid conditions and assumptions as those used to determine the nodal prices used for imbalance/spot energy settlements and transmission usage charges.

FTR Obligations and Options

The FTRs offered in PJM and New York and similar rights proposed for New England function as "obligations." The defining feature of an obligation FTR is that the holder of the FTR is not only entitled to receive the difference in nodal prices if the difference is positive, but is obligated to pay (in the RTO settlements) the difference in nodal prices if the difference is negative. This would be the case if the nodal price at the point of receipt defined by the FTR were higher than the nodal price at the point of delivery defined by the FTR.

Obligation FTRs can be used to provide hedges for proposed counter-flows that help to relieve congestion. Counter-flows reduce the marginal cost of any congestion redispatch. In cases where the locational price differences are negative, a party with an obligation FTR would either provide the counterflow (in which case it would receive through the RTO settlements the difference in nodal prices to offset its FTR obligation payments) or pay the marginal redispatch costs (defined as the difference in nodal prices) for the FTR's locations.

An obligation FTR provides a means to sell congestion management services on a forward basis. Parties willing to undertake the obligation for those locations likely to require payment can therefore be paid to do so. Willing parties are essentially betting that the amount they are paid in advance to take on the obligation will turn out to be greater than the cost of either providing the counter-flows (redispatch) in real time or paying for redispatch. The obligation FTRs that turn out to be "negative" FTRs simply become another product that parties can acquire and trade to support their commercial objectives.

When parties commit either to provide counter-flow/redispatch or to pay for it, these binding commitments allow more transactions to flow in the opposite direction, expanding use of the grid for the entire market. Obligation FTRs thus allow an RTO to maximize grid use, accommodate more transactions and issue more rights.

FTRs can also be defined as "options." When congestion arises and results in nodal price differences, the option FTR holder is hedged against the congestion/usage charges when the nodal price difference is positive, but the option holder is not obligated to pay the nodal price difference when this difference is negative. Not having this obligation will be valuable to some participants, particularly those who do not plan to implement corresponding (counter-flow) schedules. Options may be less important to those prepared to provide the counter-flow or willing to pay for it (pay the marginal redispatch costs). Moreover, if all FTRs are options, no party is obligated either to provide or to pay for counter-flows through the rights system. Without these commitments, fewer transactions can flow, and the RTO may have to allocate fewer rights if they are all options.

In issuing FTRs, an RTO would use a simultaneous feasibility test, which ensures that the total rights issued can be accommodated on the grid under expected conditions. If the issued rights meet this test, then the RTO would also know that as long as the grid has the capacity assumed under these conditions, it will always collect enough congestion revenues (from nodal prices and usage charges) to fund all of the FTRs, even if the dispatch in each settlement period is different from the dispatch assumed in the simultaneous feasibility test.[47] That is, congestion revenues collected by the RTO will be adequate to cover FTR payments owed by the RTO.

The linkage between the simultaneous feasibility test and the FTR revenue adequacy condition is an important factor in preserving the quality and value of the FTR hedges, and thus in defining how many FTRs to issue. If the test is not met, revenues may often be inadequate to fund the FTRs, so their hedging value will be diminished. Applying this test is relatively easy if the FTRs are obligations. Until now, however, there has been no workable method for applying the simultaneous feasibility test to FTRs defined as options. This was the primary reason why PJM and New York selected obligations at the time they initiated FTRs. If this problem can be solved, and a method defined to determine how much of each type of right to allocate, there need be no impediment to defining both obligation and option FTRs.[48] In theory, FTRs could be defined as either obligations or options, and the RTO could issue some of each, depending on the market's preferences.

Even when the simultaneous feasibility test is met, FTR revenue adequacy is not assured when grid conditions change. For example, a line outage that was not accounted for when the FTRs were issued can prevent certain schedules from being implemented, and thus reduce the congestion revenues collected by the RTO. This can result in the RTO having insufficient congestion revenues to fund all of the outstanding FTRs. When revenue inadequacy occurs, the RTO must have some rule to define how to deal with this shortfall.

Several mechanisms are feasible. First, the RTO will collect excess congestion revenue during some settlement periods, and it can use those surpluses to fund revenue shortfalls in other periods. The carryover period might extend a month, six months, or longer. Residual shortfalls could then be handled through a *pro rata* funding of the FTRs. Alternatively, in addition to using surpluses carried over from previous periods, the RTO rules can provide that it will fully fund the FTRs in each settlement period and add any residual deficit or surpluses (after applying surpluses from other periods) to the revenue requirements for the transmission tariff.

Under a third approach, the RTO and its associated grid owner(s) can develop a performance incentive that rewards the grid owner(s) for proper grid maintenance that minimizes grid outages, while holding the owner(s) responsible for FTR revenue shortfalls attributed to grid outages.

[47] Scott M. Harvey, Susan L. Pope, and William Hogan, "Transmission Capacity Reservations and Transmission Congestion Contracts," June 6, 1996.

[48] In approving the redesign of New England's congestion management markets to implement nodal pricing, FERC directed the ISO-NE to offer both obligation and option FTRs. Work to test a possible approach to the simultaneous feasibility test for FTR options is underway but not yet completed.

FTR funding is thus assured, and the mechanism can encourage efficient grid maintenance through performance incentives.[49]

Financial Transmission Rights and Market-Driven Transmission Investments

An important attribute of FTRs is that they are tradable property rights that reflect the market value of using the grid. They can be acquired in advance through RTO allocations and auctions, and once acquired, they can be traded in forward secondary markets. The prices at which parties acquire and trade FTRs in advance indicate the market's forward valuation of grid usage. Parties will tend to acquire and trade FTRs at prices they think represent the expected value of the stream of payments the holder will receive over the FTR period. The prices will tend to reflect the expected value that the market places on avoiding congestion charges. FTRs thus provide a measure of the market value of upgrading the grid to relieve the same congestion.

Because forward FTR prices can signal the market's economic value of grid improvements, the ability to acquire FTRs can be a mechanism to support market-driven investments in grid upgrades. The award of FTRs thus provides an opportunity for investors to receive the market value of their grid upgrades. An important rule in the Standard Market Design would therefore award to those who fund grid upgrades the incremental FTRs that become possible as a result of those upgrades.[50]

- An RTO should award to those who invest in grid upgrades the incremental FTRs that the upgrades make possible under the simultaneous feasibility test. The length of the FTRs should be consistent with the longevity of the investment, or in any event, long enough to allow a reasonable period for recovery of the cost of the investment and profit expectations.

- An RTO should have a defined method for determining the incremental FTRs made possible by a proposed upgrade.

These principles will allow and encourage market-driven investments in transmission. Many such investments may be possible, and over time they may become an important mechanism for funding transmission expansions.[51] However, it can be expected that market failures will prevent some economically justified transmission investments from being pursued in response to market price incentives. For example, economies of scale may warrant investments greater than can be supported by the market-driven investment at any given moment, and free-rider problems will arise where upgrades provide benefits to many parties, making the formation of voluntary investment coalitions somewhat difficult.

Some form of regulatory backstop is therefore needed for an indefinite period to ensure that appropriate grid investments occur in the face of market failures. The investments would be subject to recovery under traditional transmission tariffs. However, regulated investments should

[49] Commonwealth Edison proposed such an incentive mechanism in its proposal for an Independent Transmission Company. See, Attachment 1 to "Petition for Declaratory Order," filed December 13, 1999.

[50] FTRs would be awarded to any entity that funded the investments, whether the original grid owners, a newly formed transco or independent transmission company, a merchant developer or a merely a group of independent investors.

[51] This means that the fixed costs of the grid will very gradually move from a regulated rate structure to a market-based pricing structure.

not be undertaken unless they can be economically justified by benefits defined in relation to expected changes in nodal prices and related usage charges, and only when the regulator can determine that one or more market failures make it unlikely the market will undertake the necessary upgrade investments.

Additional Features Compatible with the Standard Design

This paper describes the essential foundation principles for a Standard Market Design. These are the minimum requirements. However, both existing and proposed markets include additional features that have been requested by market participants and proven helpful in making the markets more workable. Without attempting to describe all desirable features, the following common features can be provided by the RTO and be consistent with the Standard Design.

- *Day-Ahead and/or Hour-Ahead Markets*. In PJM, New York and New England, participants have generally supported having the ISO coordinate bid-based short-run forward markets. Participation in these forward markets is more or less voluntary. The markets are based on participant offers and bids and arranged around an ISO-coordinated security-constrained economic dispatch. The ISO uses the price offers and bids to clear each market and define market-clearing prices for energy bought and sold in each market. Prices are defined using locational marginal pricing, so the pricing reflects the marginal cost of any redispatch needed to accommodate the forward market schedules. Parties that commit to buy and sell energy in the forward market are held financially responsible for these purchases and sales and settled in that market at the forward market-clearing prices. In the real-time market settlements, deviations from the quantities committed to in forward market are then settled at the real-time market prices.

- *Unit Commitment Service*. In PJM, New York and New England (proposed), the ISO offers a voluntary unit commitment service, based on three-part bids, which indicate each plant's incremental energy costs, start-up costs and minimum generation costs. Generators are allowed to self-schedule their own units, but they may also allow the ISO to determine the most economic unit commitment for their plants. Participating generators are guaranteed recovery of their start-up and minimum generation costs in the event they fail to recover these costs from the prices received in the ISO-coordinated markets.

Mitigation of Market Power

The potential for market power will need to be addressed in each market region. This is a complex topic well beyond the scope of this paper. In general, the exercise of market power is more likely and can be more profitable under a defective market design; hence close adherence to the principles in the Standard Market Design is an essential strategy for minimizing the risks of market power. While it may be easier to deal with market power issues prior to market commencement, it is likely that grid constraints will still leave pockets in which strategic bidding or withholding to raise prices above efficient levels can be successful if not mitigated. These concerns can be addressed through mandatory availability requirements, forward contracts for "reliability-must-run" units, and selective bid caps on generators at locations where grid constraints allow the exercise of market power. The availability requirements ensure that units essential for reliability are offered to the RTO through bids into its real-time markets. The bid caps serve to limit the prices that can be offered by those in strategic grid locations where, because of grid constraints, little or no effective competition exists. The Standard Market Design simplifies market power mitigation by allowing mitigation rules that would be consistent with a competitive market.

Related Issues

The principles described in this paper provide much of the essential foundation for a workable, competitive electricity market. However, there are many other design details that will affect how well a market will function. An especially important topic not discussed in this paper concerns the method by which market prices are defined and possibly limited during periods of supply shortages, how demand bids can be used to define prices, and how prices are set during real shortage conditions in the absence of demand-side responses. A related topic concerns the merits of capacity requirements and the possible use of capacity markets or other mechanisms to help assure supply adequacy. These are extremely complex topics for which one or more separate papers would be appropriate. Additional papers would discuss such topics as state regulatory actions to support markets, the complementary role of retail pricing and the conditions for enhancing demand-side responses to market prices. These issues can be examined intelligently within a framework in which the core principles described in this paper are accepted as the foundation. Without the foundation provided by the Standard Market Design, discussion of these and other obviously important topics would have much less value.

Conclusion

FERC's goals for efficient competitive markets would be substantially furthered if RTOs were strongly encouraged to adhere to the foundation principles in the Standard Market Design. The principles define RTO market support functions and pricing methods that have a proven record of success in the Eastern United States and elsewhere. The Standard Design is based on a bid-based, security-constrained economic dispatch, coordinated by the RTO, with the market prices used for settlement defined by locational marginal costs. Financial Transmission Rights that hedge locational price differences and can be traded and acquired in advance to provide price certainty complete the basic design and provide a basis for market-driven grid investments and trading of risk management instruments. Within this Standard Design framework, participants can freely engage in bilateral trading, self scheduling and other decentralized trading or rely on coordinated spot trading settled at market-clearing prices defined by the RTO. Trading can then be simplified and made more liquid by defining financial trading hubs or other nodal pricing aggregations for settlements.

Experience with markets based on these principles is sufficient to establish them as the benchmark for RTO functions and "best practice" market designs. Further innovation and experiments with design principles can be considered, but proponents should bear the burden of showing that untried approaches are likely to work, pose acceptable risks and are more likely to achieve the goals of an efficient competitive market than the Standard Design.

Panel Session 3

Alternative Wholesale Power Market Structures for California

External Affairs

Philippe Auclair
Manager, Regulatory Affairs

California Energy Commission Workshop on Wholesale Market Designs
November 7, 2001

Mirant Statement

- Mirant is pleased to participate in this valuable workshop to assist the State of California in finding a solution for its dysfunctional wholesale market design. Today's agenda and list of panel participants provide us with the comfort that the problems confronting the state have been identified and solutions are being sought.
- However, also know we have over 200 employees in California who worked 24 hours, seven days week to help keep the lights on during California's energy crisis. The assertion that withholding energy was a basis in California for driving prices higher is unfounded. Mirant has always operated fairly, honestly and transparently in the energy marketplace.

Wholesale Electricity Market Design

MIRANT

- **A workably competitive electricity wholesale market design (the financial model) must meet three objectives:**
  - **Objective 1**: The financial model must reflect the complex nature of the electricity network and must not succumb to convenient financial fictions.
  - **Objective 2**: The financial model must be transparent and must be easily understood by policy makers, regulators and market participants.
  - **Objective 3**: The financial model must recognize the basic principle underlying commodity markets. **"A complete wholesale commodity market comprises both spot markets and contract markets. Each is a complement of the other."**

Wholesale Electricity Market Design

MIRANT

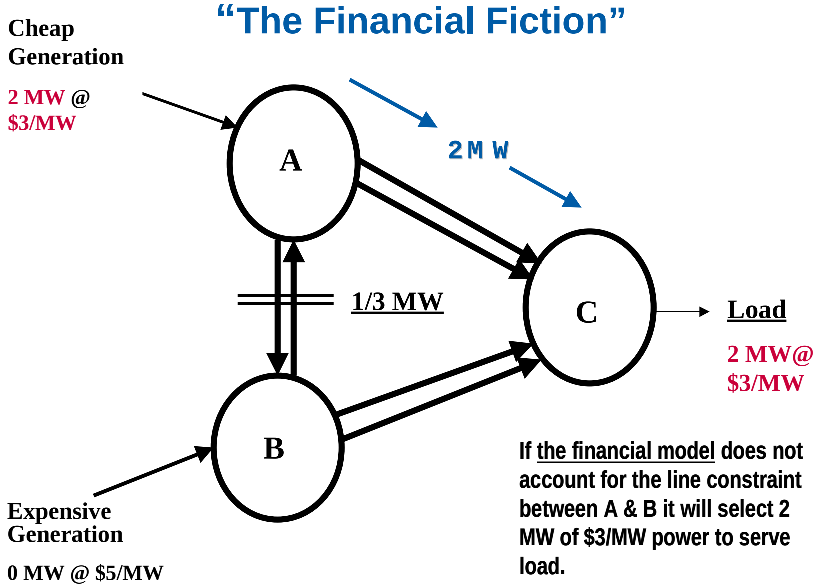
- **Why is objective 1 important?**
  - **If the electricity wholesale market design does not accurately reflect the temporal and spatial nature of the electricity network, then:**
    - **The market price signal does not provide proper instructions for a feasible short-run generation dispatch.**
    - **The RTO is forced to rely on unpredictable, opaque administrative mechanisms to instruct generation dispatch.**

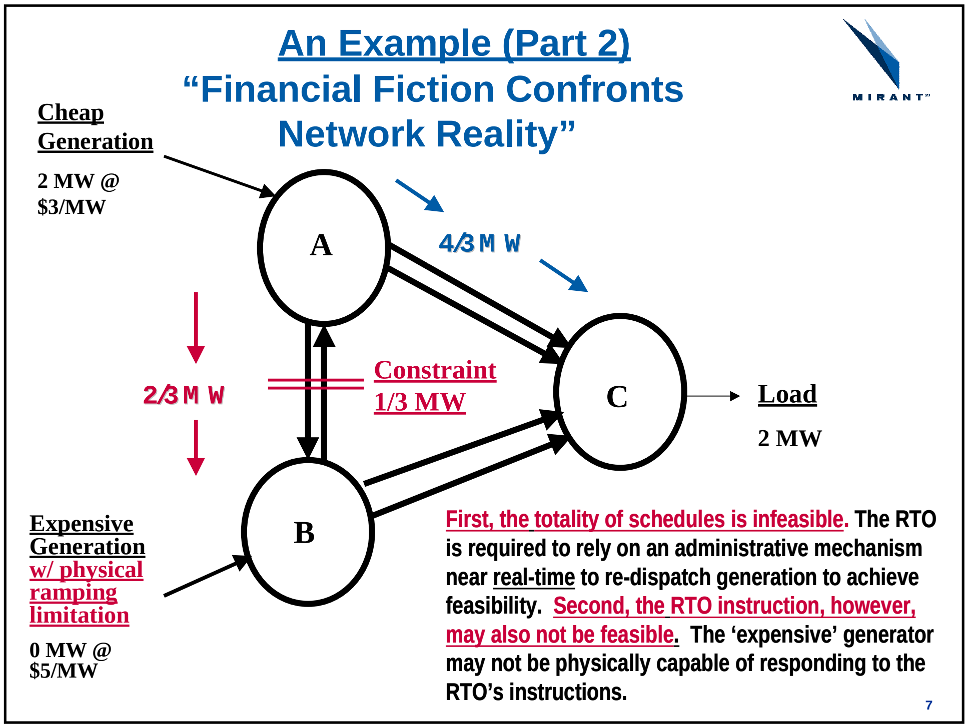
Wholesale Electricity Market Design

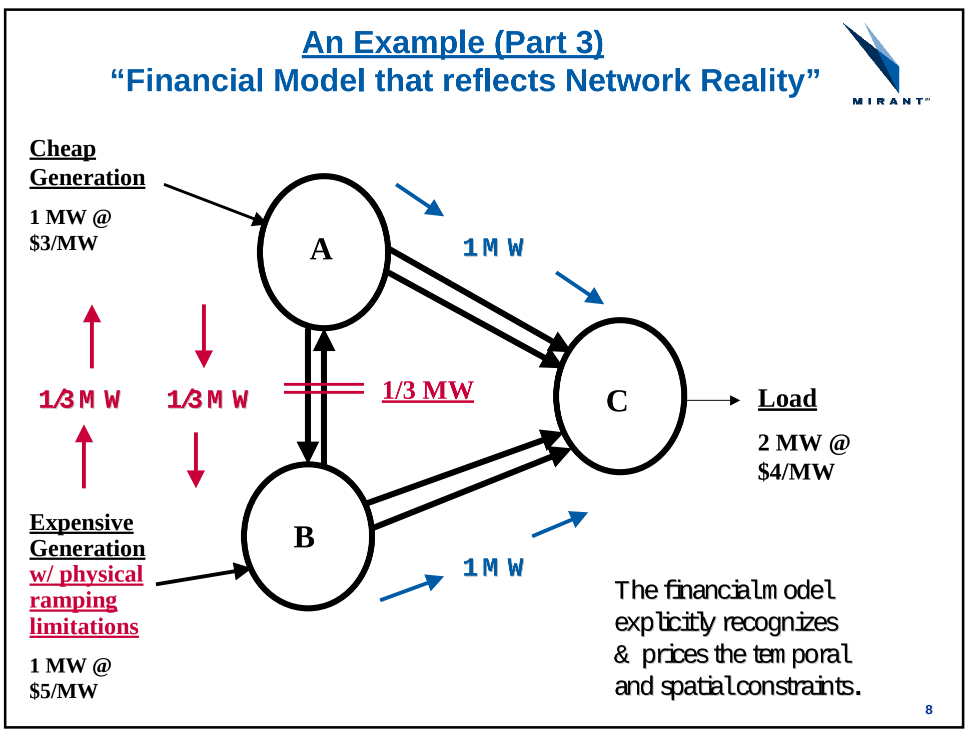
- **Why is objective 1 important? (Continued)**
  - Administrative rules lead to unintended consequences that create the need for more rules that produce more unintended consequences that create the need for more rules. This spiraling effect creates uncertainty that in turn add to costs.
  - Market participants, regulators and policy makers become confused with this "random" mix of market & administrative approaches to dispatch generation and to financially settle the dispatch.
  - **What is seen as a failure in the market mechanism is actually a failure in an 'ad hoc' administrative process that has usurped the role of the market mechanism.**

An Example (Part 1)

"The Financial Fiction"




An Example (Part 2)
"Financial Fiction Confronts Network Reality"
MIRANT
Cheap Generation
2 MW @ $3/MW
A
4/3 MW
2/3 MW
Constraint 1/3 MW
C
Load
2 MW
B
Expensive Generation w/ physical ramping limitation
0 MW @ $5/MW
First, the totality of schedules is infeasible. The RTO is required to rely on an administrative mechanism near real-time to re-dispatch generation to achieve feasibility. Second, the RTO instruction, however, may also not be feasible. The 'expensive' generator may not be physically capable of responding to the RTO's instructions.
7


An Example (Part 3)
"Financial Model that reflects Network Reality"
MIRANT
Cheap Generation
1 MW @ $3/MW
A
1 MW
1/3 MW
1/3 MW
1/3 MW
C
Load
2 MW @ $4/MW
B
1 MW
Expensive Generation w/ physical ramping limitations
1 MW @ $5/MW
The financial model explicitly recognizes & prices the temporal and spatial constraints.
8

Wholesale Electricity Market Design

- **Why is objective 1 important? (Continued)**
  - **The existing transmission and generation infrastructure is not only inefficiently used, but also underutilized.**
  - **The market price signal will not provide useful information for long-run investment decisions. Administrative and regulatory mechanism necessarily follow.**

Wholesale Electricity Market Design

- **Why is objective 2 important?**
  - **If the electricity wholesale market design is not transparent, not consistent, then policy makers, regulators and market participants 'unfairly' lose faith in the electricity market mechanism as a tool for dispatch decisions and investment activity.**
  - **What is seen as a failure in the market mechanism is actually a failure in an 'ad hoc' administrative process that has usurped the role of the market mechanism.**
  - **Ironically, the push to re-regulation invites, as a solution, the very 'ad hoc' administrative processes that have failed the industry.**

Wholesale Electricity Market Design

- Why is objective 3 important?
  - If the electricity wholesale market design only includes spot (day-ahead & real-time) markets, then:
    - All market participants (both suppliers and consumers) are faced with unnecessary risk and lack of flexibility to make efficient decisions based on prices and quality of service.

Wholesale Electricity Market Design

- Why is objective 3 important? (Continued)
  - If the electricity wholesale market design only includes forward contract markets, then:
    - All market participants (both suppliers and consumers) are faced with unnecessary risk and lack of flexibility to make efficient decisions based on prices and quality of service.

Wholesale Electricity Market Design

- **How can Objective 1 be met?**
  - **Design a transparent financial model that explicitly accounts for and prices on a day-ahead & real-time basis all temporal and spatial constraints that are inherent in an electricity network.**
  - **Under such a design, market prices will contain the necessary information to instruct a feasible as well as a least cost short-run dispatch.**
  - **It is not necessary to reinvent the wheel. A workable wholesale electricity market design already exists:**
    - **A bid-based 'physically feasible' least-cost auction is probably needed for the RTO's real-time market. The mechanism to implement this auction must reflect the realities of the electricity network.**

Wholesale Electricity Market Design

- **How can Objective 1 be met? (Continued)**
    - **Day-ahead markets in energy and capacity reserves must be implemented.**
    - **The day-ahead markets in energy and reserves must explicitly recognize and price all temporal and spatial constraints.**
    - **The design of the day-ahead market must conform to the design of the real-time market.**
    - **The day-ahead market must be reconciled with the same network model employed by the RTO for its real-time market.**
    - **The RTO network model must be made available to all market participants.**

Wholesale Electricity Market Design

- **How can Objective 2 be met?**
  - The day-ahead market and real-time market must be coordinated closely so that the transition from one to the other is seamless.
  - Simplify the institutional and regulatory structure that administers the electricity markets.
  - Given the complexity of the electricity market product, simplicity can only be feasibly obtained by reducing the complexity of the institutional and regulatory structure.
  - Having more than one RTO and more than one market monitor will not only create inefficiencies but put pressure on policy makers and regulators to re-regulate the industry.

Wholesale Electricity Market Design

- **How can Objective 3 be met?**
  - First, establish a transparent and internally consistent spot market. To quote Larry Ruff in "Stop Wheeling and Start Dealing: Resolving the Transmission Dilemma", Electricity Journal, June 1994, page 27.
    - "Paradoxically, the most important reason to create open, dispatch-based spot energy markets is not to improve short-run system operations, but to facilitate competition in the long-run markets for contracts."
  - Second, allow liquid competitive contract markets (bilateral and standardized contracts) to develop around open and transparent spot markets. Only then, will competitive electricity markets be given the chance to be fairly judged.

Structuring a Wholesale Electricity Market

Prof. Charles D. Kolstad
University of California at Santa Barbara
Bren School of Environmental Science & Management

Market Goals

- Competitive retail prices
- Stable retail prices
- High reliability
- Responsiveness to growth
- Incentives for long-term capital investment
- Only minimal oversight necessary
- Financially healthy firms

Lessons from the past year

- Cannot rely entirely on spot market
- Spot Market in Cal had two different problems
  - Price-manipulation too easy
  - Price spikes too high
- Nearly all commodity markets rely on a mix of
  - Spot purchases
  - Long-term contracts
- Long term contracts absent from Cal. Mkt
  - Incentives must exist for long-term contracting
  - Vesting contracts would have been a good idea
- Failure to deregulate retail largely responsible for where we are

First: Choose Retail Structure

- Reregulated utilities
- Retail regulation with wholesale deregulation
- Retail competition with utilities selling distribution and transmission services

Bren School of Environmental Science & Management, UCSB

Re-regulated Utilities

- Business-as-was-usual
- Utilities re-contract for long-term and short-term resources
- Oversight and rate-of-return regulation

Retail Regulation with Wholesale Deregulation

- Need mix of term contracts
- Need some long term contracts
  - Requirement that x% of expected demand be satisfied with term contracts (mixed terms)
  - Rely on Power Authority
- Doesn't matter who holds long-term contracts
  - State of California
  - Utilities
- Cannot rely on utilities to contract unless
  - Oversight of contracting
  - Stipulation of mix of contracts

Bren School of Environmental Science & Management, UCSB

Retail Deregulation

- Utilities provide transmission/distribution services
  - Rates regulated
  - Separate from power/energy charges
- Model works well in Australia, UK & Norway
- Appropriate mix of price stability, price competitiveness and reliability will emerge from market
- Need oversight of retail product offerings to assure reliability and price stability

Needed: Antitrust Guidelines for Electricity Markets

- Need equivalent of "Merger Guidelines"
- Concentration cannot be such that price manipulation is possible
- Better understanding of what constitutes suspect structure
- Better understanding of what constitutes suspect conduct

Spot vs. Term

- Limited Bilateral Mkt w/ robust spot mkt
  - What we have had
- Robust Bilateral Mkt w/ limited spot mkt
  - Desirable but will not automatically emerge
- Mixed market with government participants
  - Govt not necessary for market to function
  - Current state contracts may be held by govt or utilities
- State investment in capacity
  - Contracts are substitutes for ownership of capacity
  - State can assure adequate long-term supply

Conclusions

- Wholesale market design depends on retail design
- Incentives necessary to provide for long-term contracting
- State of California may or may not play a role as provider of term power

CALIFORNIA ISO

California Independent
System Operator

Designing a Wholesale Market Structure for California's Electric Industry

Lorenzo Kristov, Ph. D
Manager of Market Design

CEC Workshop
November 7, 2001, Sacramento

CALIFORNIA ISO

California Independent
System Operator

Overview

- **Some Basic Design Principles for a Restructured, Workably Competitive Electric Industry**
- **Who Needs Spot Markets?**

Disclaimer -- This presentation contains the opinions of the author, not necessarily the policies and positions of the California ISO.

CALIFORNIA ISO

California Independent System Operator

Some Basic Design Principles

- Articulate all policy objectives of industry restructuring.
  - Enable customer choice?
  - Create efficient consumption incentives?
  - Attract market-based investment in new generating capacity?
  - Create incentives for transmission investment?
  - Stimulate new technology (generation <-> consumption)?
  - Pay off high-cost utility investments?
  - Maximize sale value of divested utility assets?
  - Maintain geographic uniformity of retail prices?
  - Stimulate supply diversity (renewables, distributed gen.)?

CALIFORNIA ISO

California Independent System Operator

Basic Design Principles (2)

- Identify competitive or potentially competitive areas within the electricity supply chain and across the customer population.
  - Fuel supply
  - Generation operation
  - Generation capacity construction
  - Transmission/distribution construction
    - T/D operation probably not competitive
  - Retail supply & associated customer services
    - Competition for small customers? Default customers?
    - Customer-side technology innovation

CALIFORNIA ISO

California Independent System Operator

Basic Design Principles (3)

- Chart causal linkages between competitive areas and policy objectives.
  - Structure of competition must lead logically to objectives.
  - Some objectives are mutually antagonistic.
- Competition in generation is not the same as -- and may be in conflict with -- competition to build new capacity.
  - Competition in generation requires excess capacity -- robust competition may be too risky to attract new investment.
  - Capacity investment requires stable, predictable revenue stream.

CALIFORNIA ISO

California Independent System Operator

Basic Design Principles (4)

- Design market and regulatory structures that provide stable linkages between competitive areas and policy objectives.
  - Customer choice -- Design must strive for cost and complexity of switching providers to be less than the benefits.
  - Generation investment -- Regulatory and market design stability; non-discriminatory, transparent dispatch procedures.
- Electric industry is permeated with externalities.
  - Society-wide -- reliable electricity is a necessity for households and a healthy economy
  - Industry-specific -- network phenomena; real-time balance.

CALIFORNIA ISO

California Independent System Operator

Basic Design Principles (5)

- The key relationship to get right is between the customer and the retail supplier (even without retail competition).
  - Retailer must know what it costs to serve the customer, i.e., how much the customer consumes when.
  - Retailer must be responsible for adequate supply to serve load.
    - One retailer's supply shortfall affects the whole system.
- Market decisions must support real-time reliable grid operation.
  - Clear boundary between decentralized (market) decisions and centralized dispatch control.
  - Transparent, predictable rules of dispatch, including out-of-merit.
  - Clear, effective incentives for following schedule & instructions.
  - Is there a necessary link between real-time dispatch and pricing?

CALIFORNIA ISO

California Independent System Operator

Who Needs Spot Markets?

- No market design is perfect; all are exploitable; complex designs more easily exploitable than simple ones.
- Theoretical market efficiency is compromised when:
  - Commodity is too non-homogeneous (temporal, geographic, and quality features are essential to consumers)
  - Information is asymmetric (e.g., sellers have better info than buyers).
  - Network externalities can not be internalized.
- Spot pricing may not be needed.
  - No necessary link between dispatch and pricing.
  - All services can be procured via bilateral contracts.